THE UNIVERSITY
of EDINBURGH

On the Energy Efficiency of Spatial Modulation Concepts

Athanasios Stavridis

A thesis submitted for the degree of Doctor of Philosophy.
The University of Edinburgh.
October 2015

Abstract

Spatial Modulation (SM) is a Multiple-Input Multiple-Output (MIMO) transmission technique which realizes low complexity implementations in wireless communication systems. Due the transmission principle of SM, only one Radio Frequency (RF) chain is required in the transmitter. Therefore, the complexity of the transmitter is lower compared to the complexity of traditional MIMO schemes, such as Spatial MultipleXing (SMX). In addition, because of the single RF chain configuration of SM, only one Power Amplifier (PA) is required in the transmitter. Hence, SM has the potential to exhibit significant Energy Efficiency (EE) benefits. At the receiver side, due to the SM transmission mechanism, detection is conducted using a low complexity (single stream) Maximum Likelihood (ML) detector. However, despite the use of a single stream detector, SM achieves a multiplexing gain.

A point-to-point closed-loop variant of SM is receive space modulation. In receive space modulation, the concept of SM is extended at the receiver side, using linear precoding with Channel State Information at the Transmitter (CSIT). Even though receive space modulation does not preserve the single RF chain configuration of SM, due to the deployed linear precoding, it can be efficiently incorporated in a Space Division Multiple Access (SDMA) or in a Virtual Multiple-Input Multiple-Output (VMIMO) architecture.

Inspired by the potentials of SM, the objectives of this thesis are the evaluation of the EE of SM and its extension in different forms of MIMO communication. In particular, a realistic power model for the power consumption of a Base Station (BS) is deployed in order to assess the EE of SM in terms of Mbps/J. By taking into account the whole power supply of a BS and considering a Time Division Multiple Access (TDMA) multiple access scheme, it is shown that SM is significantly more energy efficient compared to the traditional MIMO techniques. In the considered system setup, it is shown that SM is up to 67% more energy efficient compared to the benchmark systems. In addition, the concept of space modulation is researched at the receiver side. Specifically, based on the union bound technique, a framework for the evaluation of the Average Bit Error Probability (ABEP), diversity order, and coding gain of receive space modulation is developed. Because receive space modulation deploys linear precoding with CSIT, two new precoding methods which utilize imperfect CSIT are proposed. Furthermore, in this thesis, receive space modulation is incorporated in the broadcast channel. The derivation of

the theoretical ABEP, diversity order, and coding gain of the new broadcast scheme is provided. It is concluded that receive space modulation is able to outperform the corresponding traditional MIMO scheme. Finally, SM, receive space modulation, and relaying are combined in order to form a novel virtual MIMO architecture. It is shown that the new architecture practically eliminates or reduces the problem of the inefficient relaying of the uncoordinated virtual MIMO space modulation architectures. This is undertaken by using precoding in a novel fashion. The evaluation of the new architecture is conducted using simulation and theoretical results.

Lay Summary

In modern life, the use of wireless communication has been extensive. In particular, wireless communication is a means of working, socializing, entertainment, and of many others. But, how wireless communication is achieved? Today, the main method of wireless communication is the transmission of Radio Frequency (RF) signals in different frequency bands. Unfortunately, the available frequency bands are limited and their use is very expensive. Therefore, the efficient use of the available frequency bands is very important. Since the late 1990s, research has shown that the deployment of multiple antennas, both at the transmitter and receiver, can offer more efficient wireless communication without requiring additional frequency bands. However, the use of multiple antennas imposes significant technical challenges and a significant increase in complexity. For this reason, the proposition and study of multi-antenna communication systems which overcome these challenges is important.

A multi-antenna communication concept that practically solves or alleviates the challenges that conventional multi-antenna architectures face is Spatial Modulation (SM). Due to its operating mechanism, SM offers a low complexity system implementation both at the transmitter and receiver. Also, for the same reason, SM is able to offer significant Energy Efficiency (EE) benefits. Motivated by this, the analysis of the EE of SM and its comparison with competing multi-antenna architectures are undertaken. Considering the whole power supply of a transmitter, this research shows that SM is able to offer significant EE benefits in comparison to the State of the Art (SotA) benchmark systems.

Due to the potential of SM, since its invention in the early 2000s, a large number of variations and modifications of SM has been proposed and studied in the published research. One variant is receive space modulation. The main characteristic of receive space modulation is that it uses advanced signal processing techniques in order to establish communication between a transmitter and a receiver. For this reason, the quality of reception at the receiver is increased. In this thesis, the performance of receive space modulation is studied in detail using a theoretical analysis and Monte Carlo simulations. Also, it is shown when receive space modulation performs better than the corresponding SotA benchmark system. In addition, in this thesis, due to the deployment of advanced signal processing techniques in the transmitter, receive space modulation is incorporated in a form of wireless communication termed as the broadcast channel.

In the broadcast channel, multiple users are served concurrently using the same RF spectrum resources. Therefore, the available RF spectrum resources are better utilized. In this thesis, using a theoretical analysis and Monte Carlo simulation results, the performance of the new architecture is characterized and compared against the corresponding performance of the SotA benchmark systems. In addition, inspired by the potential of SM and receive space modulation, a new distributed communication architecture, based on SM and receive space modulation, is proposed and studied. Finally, it is shown that the new architecture outperforms the corresponding SotA systems.

Declaration of Originality

I hereby declare that the research recorded in this thesis and the thesis itself was composed and originated entirely by myself in the Institute for Digital Communication of the School of Engineering, at The University of Edinburgh.

The only exception to the above is the following:

- The MATLAB code used as a basis for the simulation setups of Chapter 3 is partially based on previous work by Dr. Hauke Holtkamp.

Athanasios Stavridis

Acknowledgements

Firstly and foremost, I would like to express my sincere thanks, appreciation, and deepest gratitude to the supervisor of my PhD studies, Professor Harald Haas. He gave me the opportunity to purse my PhD degree, which has been one of my most desirable dreams until this moment. Therefore, the volume of my gratitude for this is unlimited. During my PhD studies, Professor Haas has been the best example of academic excellence who inspired me the most. His insightful advice and technical guidance has shaped the way of my thinking in the most positive direction. Thank you Professor for everything.

I would also like to express my thanks to my second supervisor, Professor John Thompson, for all of his support during my studies. Whenever it was required, he provided me with the best advices.

I am really grateful to my research coauthors during my PhD studies, Dr. Sinan Sinanović, Dr. Marco Di Renzo, and Dr. Dushyantha Basnayaka. Each one of them helped me in the most valuable way. During my first years of study, Sinan simplified and shed light in every technical problem I had. Usually, this was happening in a daily basis. Marco has been a great inspiration and always eager to explore new problems. He provided me with technical advices that put my research in a higher level. Dushyantha really helped me in the last stage of my PhD. Our discussions enabled me to produce the best result. My research would not be the same without them. Thank you for this.

Many unlimited and special thanks are expressed to Adean Lutton for proofreading parts of this thesis. Adean made sure that, from a linguistic point of view, these parts are presented in the best way. Also, she positively affected my style of writing in an invaluable way. Thank you very much Adean.

At this point, I would like to thank my good friend and current flatmate Dominik R. Laetsch for generously offering his help and proofreading parts of this thesis. Beyond this, during the past months, Dominik provided me with some very interesting conversations in technical and non-technical issues. Thank you Dominik for being such a good friend.

Furthermore, I would like to acknowledge the financial support of the European Union to my

PhD studies. The work presented in this thesis was supported by the European Union's Seventh Program for Research, Technological Development and Demonstration under Grant PITN-GA-2010-264759 (Greenet project). In addition, I would like to thank all the people of the Greenet project for our three years fruitful collaboration.

The office hours during the past years have been uncountable. However, due to my office friends, Yannis, Dobroslav, Stefan, Rodrigo, Miltos, Ebtihal, Chunli, Nick, Boggo, Harald B., Abdelhamid, Xiping, and Myriam, I have some really nice memories. In additions to this, they were always available for short or even longer technical discussions. Thank you all. Furthermore, many special thanks go to my Edinburgh friends, Eleanna, Giorgos T., Zafiris, Ilenia, Ntina, and Konstantina. Their presence made the already beautiful Edinburgh even more beautiful. My good friends and past flatmates, Iraklis and Giorgos K., deserve an unlimited number of thanks. They provided me with support and friendship in every moment during the last years of my PhD studies, day or night.

Last but not least, I would like to express my gratitude and love to the most valuable people of my life, my family. All these years, they provided me with love and support in every possible way. I could never succeed anything without them. They really shaped my values, character and ethics. What I feel and what I owe to them cannot be expressed with minor things like words.

As I read in a source (book, web page, or magazine), which unfortunately I do not remember in order to cite, *"I buy the first round of drinks to the ones who are mentioned in this acknowledgment, and the next many rounds of drinks to the ones who are unintentionally missed to be mentioned here, but they should had been"*.

Contents

List of figures

List of tables

Acronyms

1G	First Generation
2G	Second Generation
3G	Third Generation
3GPP	3rd Generation Partnership Project
4G	Fourth Generation
5G	Fifth Generation
ABEP	Average Bit Error Probability
AF	Amplify-and-Forward
AI	Antenna Interface
AMPS	Advanced Mobile Phone Service
AS	Antenna Selection
AWCGN	Additive White Complex Gaussian Noise
BER	Bit Error Rate
BI	Baseband Interface
BLAST	Bell Laboratories Layered Space-Time
bpcu	bits per channel use
bpsp	bits per signaling period
BPSK	Binary Phase Shift Keying
BS	Base Station
CDMA	Code Division Multiple Access

CD	Centralized Detection
CF	Compress-and-Forward
CMOS	Complementary Metal Oxide Semiconductor
CoMP	Coordinated MultiPoint
CSI	Channel State Information
CSIR	Channel State Information at the Receiver
CSIT	Channel State Information at the Transmitter
DCMC	Discrete input Continues output Memoryless Channel
DC	Direct Current
DD	Distributed Detection
DF	Decode-and-Forward
DN	Destination Node
DH-DAC	Dual Hop-Distributed Alamouti Code
DH-HSM	Dual Hop-Hybrid Spatial Modulation
DH-HSSK	Dual Hop-Hybrid Space Shift Keying
DH-w1R	Dual Hop-with one Relay
DH-wBRS	Dual Hop-with Best Relay Selection
EARTH	Energy Aware Radio and neTwork tecHnology
EB	Exabyte
EBF	Eigen-BeamForming
ECG	Equal Gain Combining
EE	Energy Efficiency
EF	Estimate-and-Forward

FDMA	Frequency Division Multiple Access
FM	Frequency Modulation
FSO	Free Space Optics
Gbps	Gigabit per second
GeSM	Generalized Spatial Modulation
GSM	Global System for Mobile communications
GSSK	Generalized Space Shift Keying
HSM	Hybrid Spatial Modulation
HSSK	Hybrid Space Shift Keying
HT	Hypothesis Test
IAS	Inter-Antenna Synchronization
ICI	Inter-Channel Interference
ICT	Information and Communications Technology
I-CSIT	Imperfect-Channel State Information at the Transmitter
i.i.d.	independent and identically distributed
IMTS	Improved Mobile Telephone Service
ISI	Inter-Symbol Interference
IS-95	Interim Standard 95
Kbps	Kilobits per second
MAC	Multiple Access Channel
Mbps	Megabits per second
MIMO	Multiple-Input Multiple-Output
MISO	Multiple-Input Single-Output

ML	Maximum Likelihood
MMSE	Minimum Mean Square Error
MRC	Maximum Ratio Combining
MRT	Maximum Ratio Transmission
MS-SM	Multi-Stream Spatial Modulation
MSR-SM	Multi-Stream Receive-Spatial Modulation
MU	Multi-User
MU-MIMO	Multi-User Multiple-Input Multiple-Output
NBE	Norm Bound Error
NLOS	Non-Line of Sight
NMT	Nordic Mobile Telephone
OFDMA	Orthogonal Frequency-Division Multiple Access
OFDM	Orthogonal Frequency-Division Multiplexing
OSTBC	Orthogonal Space-Time Block Coding
PA	Power Amplifier
PAPR	Peak-to-Average-Power Ratio
P-CSIT	Perfect-Channel State Information at the Transmitter
P-CSIR	Perfect-Channel State Information at the Receiver
PDF	Probability Density Function
PEP	Pairwise Error Probability
PSK	Phase Shift Keying
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service

RCR	Relative Complexity Reduction
RAER	Relative Average Energy Reduction
RF	Radio Frequency
RN	Relay Node
R-SM	Receive-Spatial Modulation
RV	Random Variable
RF-TRX	Small-Signal RF Transceiver
SDMA	Space Division Multiple Access
SER	Symbol Error Rate
SIC	Successive Interference Cancellation
SIMO	Single-Input Multiple-Output
SISO	Single-Input Single-Output
SM	Spatial Modulation
SMX	Spatial MultipleXing
SN	Source Node
SNR	Signal-to-Noise Ratio
SOCP	Second Order Cone Programming
SotA	State-of-the-Art
SRA	Stochastic Robust Approximation
SSK	Space Shift Keying
STBC	Space-Time Block Coding
STSK	Space Time Shift Keying
SVD	Singular Value Decomposition

SU	Single-User
TACS	Total Access Communications System
TCM	Trellis Coded Modulation
TDD	Time Division Duplexing
TDMA	Time Division Multiple Access
TR	Tikhonov Regularization
V-BLAST	Vertical-Bell Laboratories Layered Space-Time
VLC	Visible Light Communication
VMIMO	Virtual Multiple-Input Multiple-Output
WCRA	Worst-Case Robust Approximation
ZF	Zero Forcing
ZMCSCG	Zero Mean Circular Symmetric Complex Gaussian

List of Principal Symbols

$\mathbf{A}$	Matrix
$\mathbf{a}$	Vector
a	Scalar
$\arg$	Argument
$(\cdot)^{-1}$	Inverse matrix
$\text{adj}(\cdot)$	Adjoint matrix
$[\mathbf{A}]_{i,i}$	(i,j) element of matrix $\mathbf{A}$
$\mathcal{C}^{n\times m}$	$n \times m$-dimensional complex space
$\mathcal{CN}(\cdot,\cdot)$	Complex Gaussian distribution
$\mathcal{CW}(\cdot,\cdot)$	Central complex Wishart distribution
$\text{Cov}[\cdot\,,\cdot\,]$	Covariance of two RVs
$\mathbf{diag}\,(a_1,\ldots,a_n)$	Diagonal matrix with a_i, being the i-th elements of its main diagonal
$\det(\cdot)$	Determinant
$\mathcal{D}(i,j)$	(i,j)-th cofactor
$\text{E}[\cdot\,]$	Expectation
$\mathbf{e}_i$	i-th column of the identity matrix
${}_pF_q\,(\alpha_1,\ldots,\alpha_p;\beta_1,\ldots,\beta_q;y)$	Generalized hypergeometric series
$\text{Gamma}(\cdot,\cdot)$	Gamma distribution
$H_0(\cdot)$	Heaviside step function
$\mathbf{I}_{n,m}$	$n \times m$ identity matrix
$\text{Im}\{\cdot\}$	Imaginary part of a complex number or matrix
$L_k^{\Delta}(\cdot)$	Laguerre polynomials
$\min$	Minimize
$\min(\cdot,\cdot)$	Minimum of between scalars
$\mathcal{N}(\cdot,\cdot)$	Real Gaussian distribution
$Q(\cdot)$	Q-function
$\text{Re}\{\cdot\}$	Real part of a complex number or matrix
$\text{tr}\,(\cdot)$	Trace
$\text{Var}[\cdot\,]$	Variance of a RV

$\mathbf{0}_{n,m}$	$n \times m$ zero matrix
$\Gamma(\cdot)$	Gamma function
$\Gamma(\cdot, \cdot)$	Incomplete gamma function
$(\cdot)^T$	Transpose
$(\cdot)^H$	Hermitian
$(\cdot)^*$	Scalar conjugate
$\otimes$	Kronecker product
$\log_2(\cdot)$	Logarithm of base 2
$\|\cdot\|_2^2$	Euclidean norm
$\|\cdot\|_F^2$	Frobenius
$\|\cdot\|$	Size of a set
$\in$	in
$(\cdot)!$	Factorial
$\binom{\cdot}{\cdot}$	Combinations
$\lfloor \cdot \rfloor$	Floor function
$\forall$	for all
$[\cdot]_{(\cdot)}$	Pochhammer symbol

List of Symbols

C_{SM}	Capacity of SM
C_1	Capacity of SM connected with the conventional signal
C_2	Capacity of SM connected with the space signal
C_{SISO}	Capacity of SISO
C_{MIMO}	Capacity of MIMO
C_{STBC}	Capacity of STBC
$\mathbf{D}$	Diagonal normalization matrix
$\breve{\mathbf{D}}$	Diagonal normalization matrix of the precoding matrix of the MIMO broadcast channel
d_{cr}	Cell radius
d_{is}	Inter-site distance
d_i	i-th element of the main diagonal of the normalization matrix $\mathbf{D}$
$d_{\text{MSR}-\text{SM}}$	Diversity order of MSR-SM
$d(\cdot)$	Number of incorrect bits between a symbol error
$\mathbf{H}$	MIMO baseband channel matrix
$\bar{\mathbf{H}}$	Known part of the MIMO baseband channel
$\widetilde{\mathbf{H}}$	Unknown part of the MIMO baseband channel
$\breve{\mathbf{H}}$	Multiuser MIMO baseband channel
$\mathbf{h}_i$	i-th column of $\mathbf{H}$
$\mathbf{h}_{\text{SIMO}}$	SIMO column channel
$\mathbf{h}_{\text{MISO}}$	MISO row channel
h	SISO scalar channel
$h_{i,j}$	(i, j) Element of $\mathbf{H}$
k_{SM}	Spectral efficiency of SM in bpcu
k_{SSK}	Spectral efficiency of SSK in bpcu
k_{GSSK}	Spectral efficiency of GSSK in bpcu
$k_{\text{R-SM}}$	Spectral efficiency of R-SM in bpcu
$k_{\text{MSR-SM}}$	Spectral efficiency of R-SM in bpcu
M	Constellation order
N_t	Number of transmit antennas

N_r	Number of receive antennas
N_s	Number of transmitting symbols streams in MSR-SM
N_a	Number of active antennas in GSSK
N_R	Number of RN
N_D	Number of antennas the DN
N_u	Number of user in a multiuser system
K_i	Number of antennas of the i-th user
$\mathcal{M}$	Discrete signal constellation
m	Slope of the load dependent power consumption
N_0	Noise spectral density per Hz
N_{RF}	Number of RF chains
$\mathbf{P}$	Unnormalized linear precoding matrix of the MIMO broadcast channel
$\mathbf{p}_i$	i-th column of $\mathbf{P}$
$\mathbf{P}_{\text{n}}$	Normalized linear precoding matrix
$\breve{\mathbf{P}}$	Unnormalized linear precoding matrix for
$P_{\text{bit}}(\cdot)$	ABEP for a given transmit SNR
$P_{\text{e}}(\cdot,\cdot)$	PEP for given symbol error and transmit SNR
$P_{\text{e}}^{i}(\cdot,\cdot)$	PEP of the i-th user
$P_{i}^{+\infty}(\cdot,\cdot)$	High SNR approximation of the PEP of the i-th user
$P_{\text{bit}}^{i}(\cdot)$	ABEP of the i-th user
$P_{\text{bit}}^{\text{System}}(\cdot)$	System ABEP
$P_{\text{CD}}(\cdot,\cdot)$	The end-to-end ABEP of DH-HSM when CD is used
$P_{\text{DD}}(\cdot,\cdot)$	The end-to-end ABEP of DH-HSM when DD is used
$P_1(\cdot)$	ABEP of the first hop of DH-HSM when CD is used
$P_2(\cdot)$	ABEP of the second hop of DH-HSM when CD is used
P_0	Minimum power consumption per RF chain when the BS is active
P	Total power transmission
P_t	Transmission power per antenna
P_{max}	Maximum transmission power of an antenna
P_{sleep}	BS power consumption when the sleep mode is applied
$P_{\text{supply}}^{\text{SM}}$	Power supply of a BS employing SM
$P_{\text{supply}}^{\text{MIMO}}$	Power supply of a BS employing conventional MIMO
P_{supply}	Power supply of a BS

R_{STBC}	Rate of a certain STBC code
R_{STBC}	STBC rate
$\mathbf{R}_T$	Transmit correlation matrix
$\mathbf{R}_R$	Receive correlation matrix
$\text{SER}_i^{+\infty}$	High SNR approximation of the symbol error rate of the i-th user
$\text{SER}_{\text{System}}^{+\infty}$	High SNR approximation of the system symbol error rate
W	Bandwidth
B	Signal bandwidth
B_c	Coherence bandwidth
$\mathbf{W}$	Linear processing matrix at the receiver
$\mathbf{W}_{\text{ZF}}$	ZF linear processing matrix at the receiver
$\mathbf{W}_{\text{MMSE}}$	MMSE linear processing matrix at the receiver
$\mathbf{w}$	Complex Gaussian vector noise
σ_w^2	Variance of scalar complex noise
$\mathbf{x}$	Transmitted baseband signal
$\mathbf{x}_i$	Transmitted baseband signal to the i-th user
$\mathbf{y}$	Received baseband signal
$\mathbf{y}_i$	Received baseband signal at the i-th user
γ	Transmit SNR
γ_1	SNR of the first hop
γ_2	SNR of the second hop
θ	Operating temperature in K
κ	Boltzmann constant
λ_i	i-th singular value of $\mathbf{H}$
$\mathbf{\Xi}$	Large scale effect of the MIMO baseband channel
$\rho_{i,j}$	Pearson product-moment correlation coefficient between two random variables

Chapter 1
Introduction

1.1 About this Thesis

Since the invention of Radio Frequency (RF) wireless communication, its spread in modern life has been constantly increasing. Especially after the introduction of the first cellular systems in 1970s, the use of wireless communication has changed from a simple voice service to more complicated and data rate demanding services, such as real time video. The main factor for this change is the developments in Information and Communications Technology (ICT) [2]. Due to these developments, the cost of wireless communication is decreased, while, the provided Quality of Service (QoS) is rapidly increased. Modern wireless communication systems are now used in an increasing number of applications, such as information retrieval from the Internet and social networking. Furthermore, in the future, wireless networks are expected to accommodate the Internet of Things, where a wide range of devices and network architectures needs to be connected [3].

Therefore, the data rate requirement of the future wireless networks becomes even more demanding [2, 4]. As demonstrated in Fig. 1.1, Cisco forecasts that in 2018, the data traffic of portable devices is expected to reach up to 15.9 Exabyte (EB) per month. As shown by Claude Elwood Shannon in [5], the most effective pathway for higher data rates is the usage of more RF spectrum. In cellular communication systems, this is achieved by spatially reusing the available RF spectrum. Indeed, throughout the different generations of cellular systems, higher data rates have been achieved by using additional RF spectrum. However, the available RF spectrum is limited and licensed. Thus, it is a very expensive and valuable resource. For this reason, any technique that offers higher data rates without the use of additional RF resources becomes increasingly desirable.

A link level technique that is able to offer higher data rates without additional RF resources is Multiple-Input Multiple-Output (MIMO) communication. Its main characteristic is the use of multiple antennas at both the transmitter and the receiver. It has been shown theoretically that the channel capacity of a MIMO system scales linearly with $m = \min(N_t, N_r)$, where N_t and

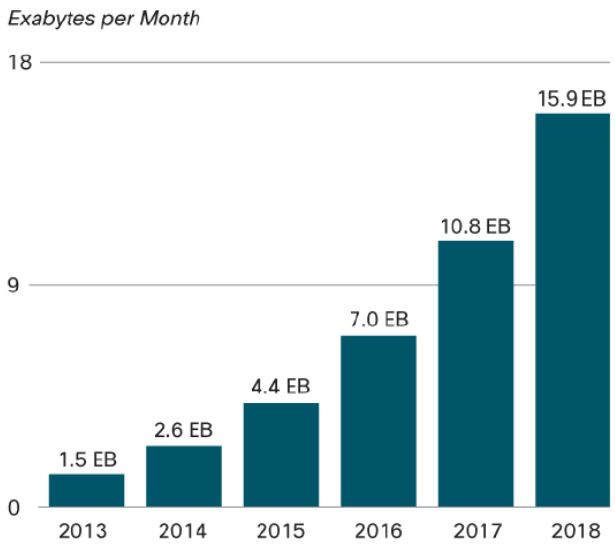


Figure 1.1: *Forecast of mobile data traffic per month, between the years 2013 and 2018, as provided by Cisco [1].*

N_r are the number of transmit and receive antennas, respectively [6, 7]. After its information-theoretical evaluation, multi-antenna communication became the main area of research in wireless communication. Since the publication of [6, 7], several MIMO architectures with different characteristics have been proposed. For example, high data rates are achieved using Spatial MultipleXing (SMX) [6]. In SMX systems, multiple parallel streams of binary information are established from the transmitter to the receiver. Also, multi-antenna communication can benefit from spatial diversity. Spatial diversity offers better Bit Error Rate (BER) performance by imposing spatial and temporal redundancy to the transmitted signal [8, 9]. Also, the deployment of multiple antennas allows the use of advanced signal processing techniques that deliver higher receive Signal-to-Noise Ratios (SNRs) [10].

Unfortunately, in a practical system, the implementation of MIMO communication imposes a number of challenges. The main challenge is the requirement of multiple RF chains in the transceiver [11]. Usually, in MIMO communication the number of RF chains equals to the number of antennas in the transmitter. Indeed, RF chains are circuits that do not follow the Moore's law. Therefore, as time passes, the efficiency of the RF chains is not expected to follow the progress of the efficiency of the rest of the circuitry. Hence, a practical implementation of a MIMO architecture may be challenging and expensive. Furthermore, a RF chain incorporates a Power Amplifier (PA) which is the main power consumer in the transmitter. Especially, this

is the case when the transmitter is a Base Station (BS) [12, 13]. Thus, a smaller number of RF chains may result in higher Energy Efficiency (EE). In addition, the computational complexity of the MIMO transceiver could become prohibitively high. For example, this is the case for SMX architectures when Maximum Likelihood (ML) detection is used. Clearly, there is a significant interest in new MIMO schemes that do not require multiple RF chains and do not significantly increase the complexity of the transceiver. Such a MIMO scheme, which does not use Channel State Information at the Transmitter (CSIT), is Spatial Modulation (SM) [14–16].

In SM, the transmitter uses its antennas in a way that only one RF chain is required. Briefly, the operating principle of SM is as follows. Provided that the impulse responses of the channels between each transmit antenna and all receive antennas are different, the bit-stream to be transmitted is divided in two portions. Note that, here, it is assumed that the number of transmit antennas is a power of two. The first portion of bits is used for the selection of the single active transmit antenna. Whereas, the second portion of bits is transmitted using a conventional symbol drawn from a standard constellation, such as Quadrature Amplitude Modulation (QAM). At the receiver side, an optimum single-stream ML detector is used in order to jointly detect the transmitted standard symbol and the index of the active antenna. In this way, the transmitted bit-stream is reconstructed at the receiver. The main advantages of SM can be summarized as: i) SM is able to achieve a multiplexing gain with the activation of a single transmit antenna; ii) SM does not require Inter-Antenna Synchronization (IAS) at the transmitter; iii) the complexity of the transmitter does not scale (significantly) with the increase of its antennas; and iv) SM offers increased EE due to its single RF chain configuration. For these reasons, SM can be considered as an alternative and promising point-to-point MIMO scheme.

During the last decade, EE in wireless cellular networks has been a popular research field due to economical and environmental reasons [17, 18]. From an operator's point of view, an efficient way to reduce its operational cost is to decrease the energy consumption of its network. In addition, the increasing evidence for the human-made climate change strongly drives the research on more energy efficient (green) solutions in wireless communication. Therefore, this thesis aims to propose a framework which quantifies the EE gains of SM compared to conventional MIMO schemes.

Inspired by the concepts of conventional SM [16] and Multi-Stream Spatial Modulation (MS-SM) [19], the authors of [20, 21] apply the concept of space modulation at the side of the receiver. In this thesis, the schemes of [20, 21] are termed as Receive-Spatial Modulation (R-SM) and

Multi-Stream Receive-Spatial Modulation (MSR-SM), respectively. Instead of activating only one antenna at the transmitter as SM, R-SM receives the transmitted signal at only one antenna at the receiver. Explicitly, the transmitter activates all of its antennas and, using linear precoding with CSIT, targets at only one receive antenna. Thus, the index of the receiving antenna can be used by the transmitter as an additional mechanism for conveying information. In a similar way, MSR-SM, using linear precoding, spatially modulates binary information to the indices of multiple receiving antennas. In an idealistic scenario of precoding, the transmitter possesses Perfect-Channel State Information at the Transmitter (P-CSIT). However, in a practical system implementation, supplying the transmitter with accurate CSIT is a difficult task. Thus, the precoding designs should be robust under imperfect or partial CSIT. In this thesis, initially, it is assumed that the transmitter is supplied with P-CSIT. Under this assumption, the statistical description of the wireless channel is used for the development of a theoretical analysis for the Average Bit Error Probability (ABEP), diversity order, and coding gain of R-SM and MSR-SM. In the next step, assuming Imperfect-Channel State Information at the Transmitter (I-CSIT), two precoding methods which deploy I-CSIT in a statistical or worst-case form are proposed. Also, the performance evaluation of R-SM and MSR-SM under channel imperfections in the transmitter is provided.

A MIMO technique that is able to deliver high data rates in a Multi-User (MU) framework without the use of additional radio resources is Space Division Multiple Access (SDMA) [22, 23]. Indeed, the broadcast channel for the downlink and the Multiple Access Channel (MAC) for the uplink serve multiple users concurrently in the same radio resources. This is done via the deployment of signal processing techniques which eliminate inter-user interference. Although SDMA schemes inspired by SM have been researched in [24, 25], the incorporation of conventional SM in a SDMA system is a challenging problem and still open. Therefore, this thesis explores the deployment of R-SM and MSR-SM in the broadcast channel.

As noted, conventional MIMO techniques encounter a number of native problems in a real system implementation. In order to overcome these problems, research has focused on Virtual Multiple-Input Multiple-Output (VMIMO) systems. The main concept behind VMIMO systems is to provide low complexity system implementations by incorporating MIMO and relay techniques. In a VMIMO system, multiple remote nodes cooperate in order to construct a transmission mechanism which virtually mimics a centralized MIMO technique. Over the recent years, there have been many proposals which combine the concepts of SM and relaying

in a virtual distributed scenario, such as [26, 27]. Inspired by the concepts of VMIMO, SM, and R-SM, the intention of this thesis is to propose a dual hop architecture which incorporates these concepts in a novel way.

1.2 Contributions

The main objective of this thesis is to study and extend the concept of SM in MIMO wireless communication systems. In more detail, this thesis focuses on: i) the energy evaluation of SM in the downlink when Time Division Multiple Access (TDMA) is the multiple access scheme; ii) the performance assessment of the extension of space modulation at the receiver side, i.e. R-SM and MSR-SM; iii) the incorporation of receive space modulation in a SDMA scheme for the downlink, i.e. the broadcast channel; and iv) the formation of a relay-based VMIMO scheme based on the concepts of SM and R-SM. The main contributions and novelties of this thesis are provided in what it follows.

In the first technical contribution of this thesis, an assessment of the EE of SM, in different types of BS, is provided. The considered types of BS are macro, micro, pico and femtocell BS. In more detail, a framework that quantifies the EE (in Mbit/J) and the power supply (in W) of SM in a BS is proposed. This framework assumes a TDMA multiple access scheme. Furthermore, it uses the Energy Aware Radio and neTwork tecHnology (EARTH) power model for the energy assessment. The EARTH power model expresses the relation between the RF power transmission of a BS and its power supply [28]. It is shown that for the same RF power transmission, the power supply of a BS which deploys SM is decreased linearly with the number of RF chains required by the same BS when the transmission technique is selected from conventional MIMO. The considered conventional MIMO techniques that serve as benchmark systems are: i) Space-Time Block Coding (STBC) [29]; ii) Multiple-Input Single-Output (MISO) with only transmit diversity [6]; and iii) MIMO [30], when the transmitter is not supplied with CSIT. The average data rate of different BSs (macro, micro, pico and femtocell BS), with respect of the total power supply, are derived using the fundamental limit of channel capacity and Monte Carlo simulations. From this study, it is concluded that for all types of BS, SM provides a range of data rates with significantly less power supply compared to the benchmark systems. In addition, using the same framework, the EE (in Mbit/J) of SM and of the benchmark systems is provided. Also, it is shown that the benchmark systems can be up to 67% less energy efficient compared to SM. It is shown that the EE of all MIMO transmission schemes (including SM) is

maximized for a certain average data rate. Note that the average data rate for which the EE is maximized is different for each considered MIMO transmission scheme.

The second technical contribution of this thesis deals with the evaluation of R-SM and MSR-SM. Initially, the theoretical analysis of the ABEP of R-SM and MSR-SM in uncorrelated Rayleigh channels is provided. This is done when P-CSIT is available and the deployed precoding method is Zero Forcing (ZF). An accurate statistical framework is used for the characterization of the received signal. Using this framework, novel and analytical bounds that characterize the ABEP of R-SM and MSR-SM are obtained. Note that, in this thesis, the complete statistical description of the wireless channel is considered without any simplification. Furthermore, the diversity order and coding gain of R-SM and MSR-SM are obtained. Monte Carlo simulation results are used for the verification of the derived bounds. The performance of R-SM and MSR-SM is compared against the performance of the corresponding conventional MIMO scheme. It is concluded that R-SM is able to offer the same or better BER performance as conventional MIMO only when the spectral efficiency is low and the number of transmit and receive antennas is low. In addition, it is shown that MSR-SM outperforms the corresponding conventional MIMO when the number of the multiple parallel spatially modulated data streams is close to the number of receive antennas. Furthermore, two regularized linear precoding methods are proposed when realistic I-CSIT is available. The first precoder is obtained using statistical I-CSIT, while, the second precoder employs worst-case I-CSIT. In addition, using an analytical framework, the instantaneous and the average power of all precoders used in this thesis are studied.

The objective of the third technical contribution of this thesis is to incorporate the concepts of R-SM and MSR-SM in a SDMA framework for the downlink, i.e. the broadcast channel. In the proposed architecture, multiple users are served from a transmitter, which can be considered as a BS. Both the BS and the multiple users possess multiple antennas. The operating principle of the proposed architecture requires that the transmitter establishes multiple parallel and non-interfering data streams to the distributed users. This is done by using linear precoding under the presence CSIT. Here, the ZF precoder is selected as the deployed precoder, due to its low complexity. The established data streams are designed in such way that binary information is transferred to the remote users using R-SM or MSR-SM. The evaluation of the new architecture is conducted using the BER metric and Monte Carlo simulations. A comparison between the performances of R-SM, MSR-SM, and the corresponding benchmark system is un-

dertaken. The benchmark system deploys conventional MIMO communication with the same linear precoding method as R-SM and MSR-SM. In order to confirm the simulation results, the theoretical ABEP of each user and of the whole system is obtained. The performance of the new architecture is also characterized by deriving its diversity order and coding gain. Note that in the theoretical analysis, the large scale channel effect of different users is taken into account. This is done in order to capture the effect of the geographical distribution of different users.

The fourth technical contribution of this thesis focuses on the proposition of a novel dual hop VMIMO architecture based on the concepts of SM and R-SM. The new distributed system, called Dual Hop-Hybrid Spatial Modulation (DH-HSM), deploys Decode-and-Forward (DF) as the relaying protocol. The aim of DH-HSM is to transfer binary information from a Source Node (SN) to a Destination Node (DN) through multiple Relay Nodes (RNs). In DH-HSM, the spatial position of the RNs is exploited for information conveyance in addition to, or even without, to the conventional way. In order to keep the complexity of the RNs and DN low, while at the same time increasing the system performance, the SN deploys linear precoding with CSIT. In this way, information is transferred from the SN to the RNs using R-SM. The RNs are able to employ a centralized coordinated or a distributed uncoordinated detection algorithm, depending on their ability to communicate. The BER metric and Monte Carlo simulation results are employed for the evaluation of DH-HSM against the: i) single relay; ii) best relay selection; and iii) distributed STBC VMIMO schemes; and iv) the direct communication link. It is demonstrated that DH-HSM offers significant SNR gains, which can be as high as 10.5 dB for a very large scale system setup. Finally, an analytical framework for the evaluation of the ABEP is developed. In this way, a verification for the presented simulation results is provided.

1.3 Thesis Outline

In the following, an outline of this thesis is given. Initially, in **Chapter 2**, a historical overview of the evolution of modern wireless communication is provided. The concept of MIMO communication is then introduced. It is explained why MIMO communication has been attracting high levels of attention from the research community. Furthermore, the categorization of MIMO architectures is outlined. Also, new trends in the research of MIMO are presented. The main limitations and challenges of MIMO communication, in a real system implementation, are summarized. The concept of SM as an alternative MIMO scheme, which alleviates or eliminates some of the previous challenges, is introduced. Finally, a brief overview of the published

research of SM and its extensions is given.

In **Chapter 3**, an energy evaluation of SM in a BS is undertaken when: i) TDMA is deployed for multi-user communication and ii) a realistic model for the wireless channels and also for the total power consumption is considered. In more detail, the wireless channels are produced using the model of [31]. Furthermore, the EARTH power model is considered. This model relates the RF power transmission to the total power supply of a BS [28]. Using the channel capacity, it is shown that, for a range of data rate, SM achieves the same data rate as conventional MIMO architectures with significantly less power supply. The EE of SM, in terms of Mbits/J, is provided. It is shown that SM can be more energy efficient than conventional MIMO system.

In **Chapter 4**, the focus is on R-SM and MSR-SM and on the theoretical evaluation of their ABEP when: i) the wireless channel is uncorrelated and follows a Rayleigh distribution; ii) the transmitter is supplied with P-CSIT; and iii) the employed linear precoding method is based on the ZF principle. As the assumption of P-CSIT is not realistic, the evaluation of R-SM and MSR-SM under I-CSIT is conducted. In addition, in this chapter, two new linear precoding methods which deploy I-CSIT are proposed. The first precoder is designed under the assumption of statistical I-CSIT. Whereas, the second precoder uses worst-case I-CSIT. Finally, analytical studies of the instantaneous and average power of all precoders deployed in this thesis are given.

In **Chapter 5**, the formation of a new SDMA scheme for the downlink is presented. The new scheme is based on the concepts of R-SM and MSR-SM. In the proposed architecture, multiple users are served using R-SM or MSR-SM. The BER performance of the new scheme is compared against the performance of the corresponding conventional MIMO scheme, revealing its limitations and potentials. Furthermore, in Chapter 5, a theoretical framework for the evaluation of the ABEP, diversity order, and coding gain of the new architecture is developed. This is done using the union bound technique. Note that this framework considers the effect of large scale fading.

In **Chapter 6**, a new and novel VMIMO architecture termed as DH-HSM is proposed and studied. In DH-HSM, communication between a SN and a RN is achieved with the cooperation of multiple RNs. This is done in two hops. During the first hop, transmission is done using R-SM, while during the next hop, SM is deployed. The main advantages and limitations of the new architecture are provided. In more detail, the BER performance of DH-HSM is compared against

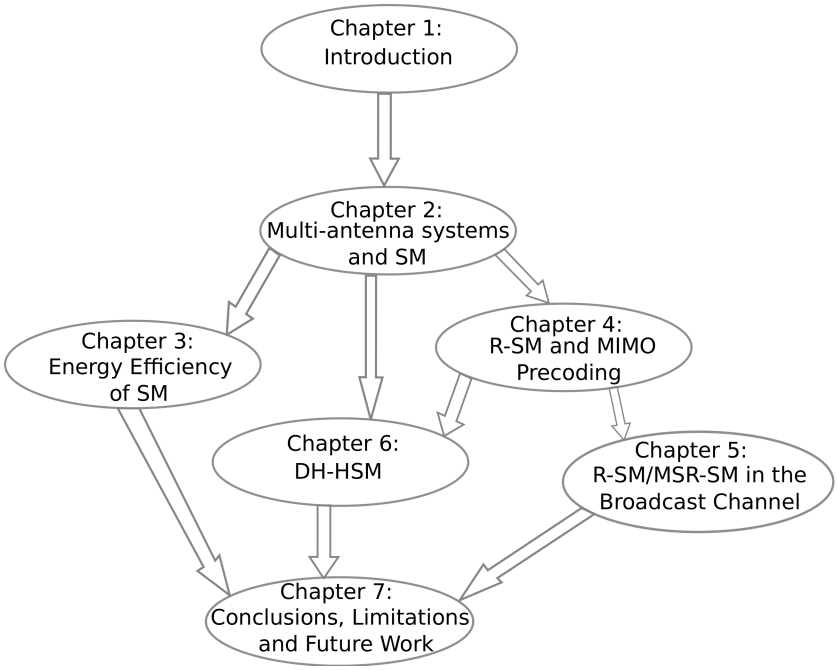


Figure 1.2: *Structure of this thesis.*

the corresponding performance of four benchmark systems. Finally, the ABEP performance of DH-HSM is studied theoretically using the union bound technique.

In **Chapter 7**, the conclusion and summary of this thesis are provided. Furthermore, the limitations and potential future extensions of this thesis are also presented.

Finally, an illustration of the structure of this thesis is presented in Fig. 1.2. Also, Fig. 1.2 demonstrates the relation among the different chapters.

1.4 Summary

This chapter provided the main motivation behind this thesis. Also, it presented a brief overview of the technical contributions and novelties of this thesis. In this context, Chapter 1 outlined the EE gains of SM. Furthermore, it briefly described the extension of the concept of SM: i) at the receiver side; ii) in a multi-user system based on R-SM and MSR-SM; and iii) in a VMIMO setup. Finally, the outline of this thesis was given.

Chapter 2
Background

2.1 The Evolution of Modern Wireless Communication

The era of modern wireless communication began in 1885 when Guglielmo Marconi demonstrated the first implementation of a radio telegraph. A few years later, in 1898, Marconi was able to establish a wireless communication via the English canal from Wimereux to Dover. The work of Marconi was enabled by the theoretical work of James Clerk Maxwell and the experimental results of Heinrich Rudolf Hertz [32]. In 1886, Maxwell predicted the existence of electromagnetic waves and described them using a mathematical theory. The experimental confirmation came a year later in 1887 from Hertz who was able to design the transmission and reception of radio pulses.

As radio technology was maturing, in 1921 the first land mobile application of the new radio wireless technology took place from the Detriot police department. Using radio transceivers (at 2 MHz) installed in police cars, a telephone system was implemented offering voice services. The advantages of this new technology became clear. However, the spread of the new technology was constrained by the channel scarcity in low frequency bands. Therefore, the wireless systems began to use higher frequency bands in order to offer more links.

An important milestone of the evolution of the wireless systems which enabled high quality in wireless communications is the concept of Frequency Modulation (FM), invented by Edwin Howard Armstrong in 1933. In 1946, Bell Systems produced a wireless system which was connected with the public telephone network. Around the same time, a similar system using FM was introduced by AT&T. This system was called the Improved Mobile Telephone Service (IMTS). Both systems were offering wireless coverage in a certain area using a fixed number of wireless channels. Today, these systems are considered to belong to the 0G pre-cellular era.

As the number of users was increasing and communication was conducted using full duplex channels, the frequency bandwidth became a valuable resource that had to be treated carefully. An effective technology that mitigated this problem and still in use today, was invented by Bell

Labs in 1947. This technology is the well known concept of cellular communication. Cellular systems are able to increase the number of served users by dividing a given geographical area into multiple cells. Each cell is allocated a portion of the available bandwidth in such way that interference between neighboring cells is avoided or reduced. The first analog cellular architecture was introduced by AT&T in 1970 and it was called Advanced Mobile Phone Service (AMPS). In this way, the era of First Generation (1G) wireless communication commenced. During the deployment of 1G, different standards were used around the globe. For example, Nordic countries, Switzerland, Netherlands, Eastern Europe, and Russia were using the Nordic Mobile Telephone (NMT) standard; the AMPS standard was adopted in North America and Australia; United Kingdom had developed the Total Access Communications System (TACS) standard; and West Germany, Portugal, and South Africa were using the C-450 standard. In general, 1G systems were able to offer a data rate ranging between 28 and 56 Kilobits per second (Kbps), and only for voice services.

The next step of the evolution of the wireless systems was the introduction of the Second Generation (2G) cellular systems. The main differences between 1G and 2G systems is that: i) 2G is solely based on digital communication; ii) 2G uses digital encryption; iii) the available bandwidth is used more efficiently; and iv) 2G offers data services, such as text and picture messages. Due to these reasons, 1G systems were replaced by 2G architectures. The multiplexing methods that 2G systems use are TDMA and Code Division Multiple Access (CDMA). The most well-known 2G wireless standards are: i) the Global System for Mobile communications (GSM), which is TDMA-based; and ii) the Interim Standard 95 (IS-95), which is CDMA-based. Despite the introduction of new standards, 2G systems are still in use in some parts of the globe. For example, various providers in Australia, Canada, and USA are planning to shutdown 2G in 2016.

The ever increasing demand for higher data rates caused the introduction of the Third Generation (3G) cellular systems. The development of 3G started since early 1980, while its commercial release was in Japan in 2001. The objective of 3G systems is to offer voice and data services. Especially for data services, the aim of 3G is to accommodate high data rate demanding applications such as mobile and fixed Internet access, video calls, and mobile television. The typical minimum data rate of 3G systems is 200 Kbps. However, later releases of 3G, often termed as 3.5G or 3.75G, offer data rates which can attain up to several Megabits per second (Mbps). The multiplexing techniques that 3G architectures use are TDMA, CDMA,

Frequency Division Multiple Access (FDMA), and Orthogonal Frequency-Division Multiple Access (OFDMA). Note that 3G incorporates the first use of multi-antenna communication.

The latest phase in the story of cellular systems is the Fourth Generation (4G) architectures. The 4G architectures offer, in addition to the existing voice and data applications of 3G, ultra-broadband Internet access. In this way, high data rate demanding applications are provided, such as gaming, high-definition television, and cloud computing. The main characteristic of 4G is its ability to offer a peak data rate of 100 Mbps in high mobility and 1 Gigabit per second (Gbps) in low mobility, using OFDMA transmission. Note that these data rates can be further increased with the deployment of multi-antenna communication. However, 4G could employ single carrier-FDMA, interleaved-FDMA, and multi-carrier CDMA. Note that 4G relies on all-Internet Protocol (IP) based communication.

As noted, the main impetus for the evolution of wireless communication systems has been the ever increasing demand for higher data rates. The theory that quantifies the theoretical upper limit of the data rate of a communication system was given by Claude Elwood Shannon in 1948, in his widely acknowledged paper entitled *"A Mathematical Theory of Communication"* [5]. Even though Shannon did not provide the way to be followed in order to achieve this theoretical limit, he provided, in addition to his limit, a quite compact theoretical framework that enables the design and theoretical evaluation of complex communication systems.

For example, the channel capacity for a simple Single-Input Single-Output (SISO) wireless channel is given as:

$$C_{\text{SISO}} = W \log_2 \left(1 + \frac{|h|^2 P}{W N_0} \right), \tag{2.1}$$

where, h represents the channel coefficient; P is the transmitted power in W; W is the available bandwidth in Hz; N_0 is the noise spectral density; and C_{SISO} is the channel capacity. From (2.1), provided that the system designer cannot control the random wireless channel and noise, the only way to increase the channel capacity is to provide additional bandwidth or to increase the transmitted power. As shown from (2.1), additional bandwidth causes a linear increase of the capacity, while higher transmitted power results in a logarithmic capacity increase. However, the parameter bandwidth is limited and usually very expensive. In addition, the unsophisticated increase of the transmitted power usually causes more problems than benefits when the SISO channel is part of a complex cellular system. This analysis can be extended to most of the modern wireless systems using similar reasoning. Due to these reasons, any candidate

solution for the future wireless networks has to consider the parameters bandwidth and transmitted power very carefully. A promising solution that might overcome the native constraints of (2.1) is multi-antenna communication. The main innovation of multi-antenna communication is the introduction of the parameter space, which offers data rate benefits without requiring any additional radio resources.

2.2 The Breakthrough of MIMO Communication

The concept of MIMO communication dates from the early 1970s. One of the first research published on MIMO communication is [33]. In addition, the concept of smart beamforming was introduced since early 1980s [34]. However, the interest in MIMO systems was prompted after the publication of [6, 7], in the late 1990s. These papers demonstrate the capacity potential of multi-antenna communication.

Since the late 1990s, the interest in MIMO wireless systems has increased exponentially. The extensive research of MIMO wireless communication systems has produced several architectures, which aim for different applications scenarios and have different characteristics. In general, MIMO systems are classified in five categories. The first category includes systems which achieve high data rate by using parallel symbol streams (spatial multiplexing). In the second category, schemes that are oriented to achieve diversity benefits and lower BER can be found. Array processing gains and higher receive SNR is the objective of the third category of MIMO. An example of this category is the linearly precoded systems using CSIT. The fourth category is composed by VMIMO systems, where multiple remote elements cooperate in order to form a distributed MIMO architecture. The fifth category includes systems which combine characteristics from the previous categories.

2.2.1 MIMO Communication: Data Rate Potential

In a single user point-to-point MIMO wireless communication system, such as the one in Fig. 2.1, a transmitter with N_t antennas communicates with a receiver with N_r antennas via a wireless channel. Provided that the wireless channel is frequency flat and slow fading, for each signaling period, the baseband representation of such communication system is given by the

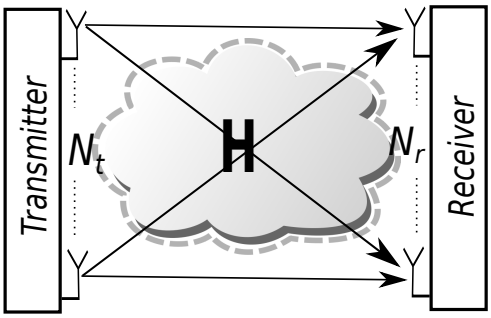


Figure 2.1: *System Model of a point-to-point MIMO communication system.*

following affine equation:

$$\begin{bmatrix} y_1 \\ \vdots \\ y_{N_r} \end{bmatrix} = \begin{bmatrix} h_{1,1} & \cdots & h_{1,N_t} \\ \vdots & \ddots & \vdots \\ h_{N_r,1} & \cdots & h_{N_r,N_t} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_{N_t} \end{bmatrix} + \begin{bmatrix} w_1 \\ \vdots \\ w_{N_r} \end{bmatrix}. \tag{2.2}$$

In a compact form, the previous equation is written as:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}. \tag{2.3}$$

In (2.3), $\mathbf{y}$ denotes the $N_r \times 1$ received signal vector. The i-th element of $\mathbf{y}$, y_i, represents the received signal at the i-th antenna of the receiver. The wireless channel is written as $\mathbf{H} \in \mathcal{C}^{N_r \times N_r}$, where its (i, j) element, $h_{i,j}$, $i = 1, \ldots, N_r$ and $j = 1, \ldots, N_t$, is the channel gain between the j-th transmitting antenna to the i-th receiving antenna. The $N_t \times 1$ transmitted signal vector is denoted as $\mathbf{x}$, with its j-th element, x_j, being the transmitted signal from the j-th transmit antenna. Finally, $\mathbf{w} \in \mathcal{C}^{N_r}$ denotes the complex vector of Additive White Complex Gaussian Noise (AWCGN) at the receiver, where $w_i \sim \mathcal{CN}\left(0, \sigma_w^2\right)$ represents the i-th element of $\mathbf{w}$, with $\sigma_w^2 = WN_0/N_r$. Note that (2.3) describes: i) a SISO system by setting the number of transmit and receive antennas equal to one ($N_t = N_r = 1$); ii) a MISO system by setting $N_t > 1$ and $N_r = 1$; and iii) a Single-Input Multiple-Output (SIMO) system by setting $N_t = 1$ and $N_r > 1$.

In practice, the wireless channel is affected by large and small scale fading. Large scale fading refers to fading due to the signal pathloss and shadowing from large objects, such as buildings. In general, large scale fading is frequency independent and changes over distances of several carrier wavelengths. Furthermore, small scale fading occurs due to the multi-path nature of the transmitted signal. In small scale fading, multiple replicas of the same signal arrive with

different delays at the receiver and have a constructive or destructive effect. Small scale fading varies over distances of the order of the carrier wavelength.

Small scale fading results in a wireless channels which is: i) frequency selective or flat and ii) slow or fast fading. More specifically, if a narrowband signal with a bandwidth, B, is transmitted over a wireless channel with a coherence bandwidth equal to B_c and it holds that, $B \ll B_c$, the wireless channel is referred as frequency flat. In this case, the fading is practically the same for whole signal bandwidth. In contrast, when it holds that, $B \gg B_c$, the channel amplitude is frequency dependent and takes different values for different frequencies. In this case, the wireless channel is referred as frequency selective. Note that a flat wireless channel does not result into Inter-Symbol Interference (ISI) at the receiver, while, a frequency selective channel imposes ISI at the receiver. In alignment with [30], the channel is considered to experience slow fading when $T_c \gg T_s$, where, T_c is the coherence time and T_s is the signaling period; when $T_c \ll T_s$, the channel is in fast fading. In this chapter, the effect of large scale fading is not considered.

The main rationale behind MIMO systems is their high spectral efficiency [6, 7]. In order to demonstrate the spectral efficiency advantages of MIMO communication over SIMO, MISO, and SISO communication, the metric of channel capacity is employed. The channel capacity of a point-to-point MIMO system with N_t transmit and N_r receive antennas, and no CSIT is given as:

$$C_{\text{MIMO}} = W \log_2 \det \left(\mathbf{I}_{N_r} + \frac{P}{N_t W N_0} \mathbf{H}\mathbf{H}^H \right) = W \sum_{i=1}^{m} \log_2 \left(1 + \frac{P}{N_t W N_0} \lambda_i^2 \right), \quad (2.4)$$

where, λ_i^2 is the i-th ordered eigenvalue of $\mathbf{H}\mathbf{H}^H$. As shown in [6, 7] and reproduced in (2.4), when the channel is full rank due to rich scattering, the capacity of a MIMO system scales approximately linear with $m = \min(N_t, N_r)$.

The channel capacity of a SIMO system is given as:

$$C_{\text{SIMO}} = W \log_2 \left(1 + \frac{P \|\mathbf{h}_{\text{SIMO}}\|_2^2}{W N_0} \right), \quad (2.5)$$

where, $\mathbf{h}_{\text{SIMO}}$ is the column SIMO channel; and the channel capacity for a MISO system is given as:

$$C_{\text{MISO}} = W \log_2 \left(1 + \frac{P \|\mathbf{h}_{\text{MISO}}\|_2^2}{W N_0} \right). \quad (2.6)$$

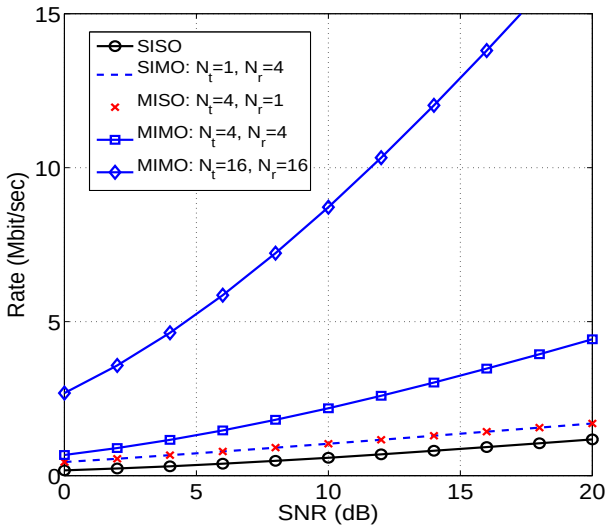


Figure 2.2: *Data rate versus SNR (dB) of SISO, SIMO* (1×4)*, MISO* (4×1)*, and MIMO* (4×4 *and* 16×16) *systems, when the available bandwidth is 200 KHz and the wireless channel follows a Rayleigh distribution.*

In (2.6), $\mathbf{h}_{\text{MISO}}$ denotes the row MISO channel vector. The capacity of a SISO system is given in (2.1).

Under the assumption that the wireless channel follows a Rayleigh distribution, with $\mathbf{H} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_r^2 \times N_r^2})$, the achieved data rate, in terms of ergodic capacity, can be evaluated using Monte Carlo simulation results. For the numerical evaluation of the ergodic capacity of a MIMO, SIMO, MISO, SISO system, (2.4), (2.5), (2.6), and (2.1), are deployed, respectively. Fig. 2.2 shows the data rate (ergodic capacity) of a: i) SISO; ii) 1×4 SIMO; iii) 4×1 MISO; iv) 4×4 MIMO; and v) 16×16 MIMO system. The utilized bandwidth is 200 KHz. As shown in Fig. 2.2, SISO communication offers 1.17 Mbps at 20 dB. This data rate is restrictive for future wireless applications, such as mobile high definition television. An improvement, which increases the data rate to 1.69 Mbps at 20 dB, is possible with the use of multiple antennas at the transmitter (MISO) or at the receiver (SIMO). However, the joint use of multiple antennas at both communicating ends provides a significant data rate potential. In more detail, a 4×4 MIMO system increases the data rate to 4.28 Mbps at 20 dB. In addition, a 16×16 MIMO system offers a significantly higher spectral efficiency, which attains to 12.02 Mbps at 14 dB. The main conclusion of this brief analysis is that MIMO communication is able to offer significantly higher spectral efficiency than SISO, SIMO, and MISO communication. Especially, as the scale of MIMO (number of transmit and receive antennas) increases, the gains are even

higher.

From an information-theoretical point of view, multi-antenna communication is a very promising and effective solution for the future wireless networks; but, MIMO is more than this. Multi-antenna communication is able to obtain multiplexing or diversity gains. Furthermore, MIMO systems can deploy advanced signal processing techniques, such as linear precoding, which deliver higher receive SNRs.

2.2.2 Multiplexing Architectures

In Section 2.2.1, the high information-theoretical data rate potential of MIMO communication is shown; but, how can this data rate be achieved by a practical system? A possible answer could be given by SMX architectures. The main principle of SMX systems relies on the creation of m parallel and independent symbol streams from the transmitter to the receiver. Usually, m is set equal to $m = \min(N_t, N_r)$. In SMX systems, every symbol period, the bitstream to be transmitted is divided into m sequences. Each sequence is encoded via a point of a conventional constellation, such as M-ary QAM or Phase Shift Keying (PSK), and transmitted through the same MIMO channel. At the receiver side, the transmitted bit-stream is reconstructed by detecting the independent parallel streams.

The concept of SMX has its origin in the early 1990s and the research published in [35]. Since then, a wide variety of SMX architectures have been proposed. The SMX architectures can be categorized based on the: i) structure of the transmitter; ii) structure of the receiver; iii) form of channel coding; and iv) the treatment of the frequency selectivity of the channel.

However, the focus on this section is on a scenario where the MIMO channel is frequency flat, the receiver possesses Perfect-Channel State Information at the Receiver (P-CSIR), and no CSIT is available. Furthermore, there is no channel coding. In this case, the receiver is able to employ either the optimum ML detector or a suboptimal one. The most important detection methods for the SMX architectures are presented in Table 2.1. The considered methods are: i) ML detection; ii) sphere decoding; iii) linear processing; and iv) Bell Laboratories Layered Space-Time (BLAST) decoding. Also, in Table 2.1, the advantages and the disadvantages of each method are summarized. A detailed discussion for each method is provided in what follows.

Initially, the focus is on the ML detector. Provided that the baseband system equation is given

Method	Advantages	Disadvantages	References
ML	Optimum performance	High computational complexity	[36]
Sphere Decoding	Near optimal performance	A trade-off between complexity and performance	[37–40]
Linear processing	Suboptimal performance	Low complexity	[41, 42]
BLAST	Better performance than linear processing	Higher complexity than linear processing	[43–47]

Table 2.1: *Detection methods for Spatial Multiplexing Architectures.*

from (2.3) and each transmit antenna emits a different symbol, the ML likelihood detector reduces to the following optimization procedure:

$$(\tilde{\mathbf{x}}) = \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2, \tag{2.7}$$

where, $\tilde{\mathbf{x}}$ is the detected symbol vector. In (2.7), each element of $\mathbf{x}$ and $\tilde{\mathbf{x}}$ belongs to the employed constellation $\mathcal{M}$. Despite (2.7) guarantees the optimum BER performance, its complexity can be prohibitively high in a real system implementation. In more detail, due to the brute force search over all of the possible transmitted vectors $\mathbf{x}$, the complexity of (2.7) follows a exponential increase in complexity [36]. Clearly, as the order M of the deployed constellation $\mathcal{M}$ and the number of transmit antennas N_t increase, the complexity scales to unacceptable levels.

Due to severity of the complexity of ML detection, several suboptimal detection methods have been developed. A solution that is able to deliver near optimal performance is sphere decoding [37–40]. The main concept behind sphere decoding is that it benefits from the statistical description of noise in order to reduce the search space of (2.7) to only a subset of all possible transmitted vectors $\mathbf{x}$. This can be undertaken by searching only through the vector symbols that lie in a hypersphere of a certain radius. The selection of the value of the radius determines a trade-off between the complexity of the receiver and the BER performance.

Linear processing is an alternative low complexity method that can be deployed at the receiver. However, this method can be applied only in system setups where the number of transmit antennas is less or equal to the number of receive antennas ($N_t \leq N_r$). This constraint guarantees that the MIMO channel matrix is not rank deficient. The baseband system representation of such a system is given as:

$$\mathbf{y} = \mathbf{W}\mathbf{H}\mathbf{x} + \mathbf{w}, \tag{2.8}$$

where, $\mathbf{W} \in \mathbf{C}^{N_t \times N_r}$ is the employed linear equalizer. All other parameters of (2.8) are defined as (2.3). Usually, the design of $\mathbf{W}$ is based on the well known ZF or Minimum Mean Square Error (MMSE) criteria [41, 42]. The ZF equalizer is written as:

$$\mathbf{W}_{\text{ZF}} = \left(\mathbf{H}^H \mathbf{H}\right)^{-1} \mathbf{H}^H, \tag{2.9}$$

while the MMSE equalizer is given as:

$$\mathbf{W}_{\text{MMSE}} = \left(\mathbf{H}^H \mathbf{H} + \frac{N_t}{\gamma} \mathbf{I}_{N_t, N_t}\right)^{-1} \mathbf{H}^H, \tag{2.10}$$

where, $\gamma = \text{E}_{\mathbf{x}} \left[\|\mathbf{x}\|_2^2\right] / \sigma_w^2$ is the transmit SNR. Usually, it is assumed that $\text{E}_{\mathbf{x}} \left[\|\mathbf{x}\|_2^2\right] = 1$.

Both equalizers of (2.9) and (2.10) have the ability to decouple the parallel symbol streams by eliminating (ZF) or reducing (MMSE) the Inter-Channel Interference (ICI). In this way, after the equalization procedure, the receiver has the option to deploy m single stream detectors (one per stream), which collectively have a significantly lower complexity. Unfortunately, linear processing at the receiver is a suboptimal technique that has poor performance, especially when ZF equalization is deployed. It is well known that ZF equalization amplifies the level of noise when the channel is in deep fade. Furthermore, MMSE outperforms ZF equalization in low SNRs, while in high SNRs, ZF approaches the performance of MMSE equalization. Due to its simplicity and low complexity, linear processing has received significant interest from industry.

Layered detection at the receiver has been proposed in order to obtain a better performance than linear processing but with a small complexity increase [43]. The introduction of layered detection in the framework of MIMO has been done in [44]. These types of algorithms are known as nulling and canceling or BLAST. They resemble to the concept of Successive Interference Cancellation (SIC) which initially was deployed in multiuser CDMA detection. In general, the operating principle of BLAST is as follows: i) for a block of symbols, during which the channel does not change, the receiver computes either the ZF or the MMSE equalizer; ii) the stream with the higher SNR is selected; ii) using the appropriate nulling vector from the linear equalizer, the selected stream is detected while the rest of the streams are treated as interference; iii) the received signal from the selected stream is subtracted; and iv) the steps between i) and iii) are repeated until all streams are detected. The error performance of each stream (layer) of BLAST is different. Furthermore, this form of algorithms experiences an error propagation between different layers. The performance analysis of Vertical-Bell Laboratories

Layered Space-Time (V-BLAST), which is a special case of layered detection, is conducted in [45]. Since the appearance of the concept of BLAST, many variations have been proposed. These variations aim to reduce its complexity [46] or to increase its performance [47].

2.2.3 Diversity Achieving Architectures

In the previous subsection, it is shown how multi-antenna communication increases the system spectral efficiency by using SMX. In addition to the SMX gains, MIMO communication is able to achieve lower BER by deploying spatial diversity techniques. However, diversity can also be used as an indirect way of increasing the data rate when it is combined with adaptive modulation.

In general, diversity techniques are categorized into two types, the large and small scale techniques. Large scale diversity or macrodiversity is connected with the phenomenon of shadowing [23, 48, 49]. In contrast, small scale techniques are connected with the small scale fading. The focus in this section is only on small scale diversity. The small scale diversity techniques are divided into the following broad subcategories: i) temporal diversity; ii) frequency diversity; and iii) spatial diversity [50]. In temporal diversity techniques, channel coding and time interleaving is used for the reduction of the effect of small scale fading. This has the disadvantage of higher delays and reduced bandwidth efficiency. Frequency diversity is a form of diversity that can be applied in wideband systems. It is obtained by transmitting the same signal via different frequency channels. Finally, spatial diversity is achieved with the introduction of spatial and temporal redundancy or the selection of the best available channel. In the first case, the same signal is transmitted (and/or received) by multiple antennas in different signaling periods, while in the latter case, the signal is transmitted and/or received by the antenna(s) with the best channel(s).

A form of spatial diversity is receive diversity, which is suitable for SIMO systems. In receive diversity systems, the receiver is equipped with multiple antennas and performs linear combining using Channel State Information at the Receiver (CSIR). The optimum method, in terms of detection or receive SNR, is Maximum Ratio Combining (MRC) [51]. A suboptimal method that belongs in the same category is Equal Gain Combining (ECG) [52].

Transmit diversity is obtained by transmitting redundant signals from the multiple antennas of the transmitter. In transmit diversity systems, the receiver can be supplied with a single or

multiple antennas. There are schemes where the transmitter is not supplied with CSIT, such as STBC [53], and schemes which benefit from the presence of CSIT, such as Maximum Ratio Transmission (MRT) [54].

Although transmit diversity exists since the early work of [55], the focus of research increased rapidly after the publication of the famous Alamouti code [8]. The Alamouti code can be applied in $2 \times N_r$ MIMO systems. It conveys two scalar symbols in two symbol periods. Its main characteristic is that it achieves full transmit and receive diversity. The structure of the Alamouti codeword is given below:

$$\mathbf{X} = \begin{bmatrix} x_1 & -x_2^* \\ x_2 & x_1^* \end{bmatrix}, \tag{2.11}$$

where, the elements of $\mathbf{X}$ belong to the employed constellation $\{x_1, x_2\} \in \mathcal{M}$. The baseband system equation of the Alamouti scheme is given as:

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{W}_n, \tag{2.12}$$

which is the extension of (2.3) in order to accommodate two symbol periods. In (2.12), $\mathbf{W}_n$ is a $N_r \times 2$ complex matrix which denotes the additive Gaussian noise over two symbol periods and $\mathbf{Y}$ is a matrix whose columns represent the received signal in two symbol periods. Therefore, in each symbol period, one column of $\mathbf{X}$ is transmitted. At the receiver side, every two symbol periods, the Alamouti codeword is reconstructed by using the following ML detector:

$$\left(\tilde{\mathbf{X}}\right) = \min_{\mathbf{X}} \|\mathbf{Y} - \mathbf{H}\mathbf{X}\|_2^2. \tag{2.13}$$

Note that, by taking into account the structure of (2.11), it can be shown that (2.13) simplifies to a form which results in linear computational complexity. For more details, the reader is referred to [8].

The Alamouti code belongs to the general category of Orthogonal Space-Time Block Coding (OSTBC). In OSTBC, the codeword design is based on orthogonal matrix designs. The main characteristic of OSTBC systems is that they achieve full diversity [53]. However, usually OSTBC experiences a rate reduction with respect to the uncoded single antenna transmission. For a complex constellation, the only code that has full rate is the Alamouti code, while for a real constellation, when the number of transmit antennas is less than 8, full rate can be ensured.

The problem of rate reduction of OSTBC can be alleviated by using non-orthogonal STBC. A type of non-orthogonal STBC that achieves full rate and transmit diversity is the diagonal algebraic STBC [56]. A form of linear constellation precoding for STBC which outperforms OSTBC is introduced in [57]. Another approach which relaxes the orthogonality constraint and offers a higher data rate than OSTBC is the quasi-orthogonal STBC [58]. In general, quasi-orthogonal STBC offers lower diversity than OSTBC. Note that many other types are reported in the literature [59–62]. In contrast to the previous types, there are codes which provide a normalized rate greater than one. For example, the linear dispersion codes are designed to offer a trade-off between spatial multiplexing and diversity gains [63]. However, as shown in [64], for all space-time wireless systems, there is a fundamental trade-off between diversity and multiplexing gains. Finally, note that the concept of space-time transmission is extended to non-coherent systems by using differential transmission [65].

An alternative concept that achieves diversity gains and allows low complexity system implementation is Antenna Selection (AS) [11]. In AS systems, a subset of the available sub-channels of a MIMO wireless channel is selected to convey information from the transmitter to the receiver. The selected subset of sub-channels (among the available ones) is the one which offer the highest receive SNR. This is done by keeping active only a portion of the available antennas at the transmitter and receiver. In the extreme case, where only one antenna is kept active at both sides of the transceiver, the diversity order is $N_t N_r$ [11]. Furthermore, if more than one antenna are kept active at the transmitter, multiple symbol streams can be established. Note that AS can be combined with other multiplexing and diversity techniques. Clearly, the number of RF chains at the transmitter and receiver equals to the number of active antennas. Given that the receiver possesses CSIR, AS can be done efficiently at the receiver. However, if the selection procedure involves the antennas of both communicating ends, the transmitter has to be informed in advanced via a feedback link.

2.2.4 Beamforming Architectures

In addition to spatial multiplexing and diversity gains, multiple antennas at the transmitter can be used for increasing the receive SNR and suppressing ICI at the receiver. In published research, these types of systems are categorized as beamforming architectures. Generally, beamforming requires CSIT. In this case, when the channel reciprocity is not satisfied, a low rate feedback link, from the receiver to the transmitter, needs to be established in order to transfer

the required CSIT. This results in a closed-loop scheme. However, in Time Division Duplexing (TDD) where the reciprocity principle can be considered as valid, CSIT can be acquired by the transmitter by deploying a training sequence transmitted from the receiver. The beamforming architectures can be classified in the following groups: i) phase shifters; ii) array processing techniques; and iii) ICI suppression techniques. Each of the previous types has its own objective.

In receive antenna arrays, where multiple antennas are accommodated, the RF signals arrive with different time delays at different antenna elements. This phenomenon causes phase shifts between the different arriving signals. If the receiver is aware of these phase shifts, it can apply phase shifters or delay elements in order to eliminate or reduce this phenomenon. Alternatively, a similar approach can be deployed at the transmitter, given that it possesses the knowledge of the occurring phase shifts. This form of processing is termed conventional beamforming [66].

In a wide range of wireless communication scenarios, the RF signal propagates only through a small number of dominant paths. In this case, either the transmitter, or the receiver, or both can adjust their beam patterns to the directions of the dominant paths. Thus, a SNR gain is obtained with respect to the omni directional beam pattern antenna arrays. This type of beamforming belongs to the array processing techniques and is extremely useful in the emerging concept of millimeter wave communication [67–69].

An additional use of beamforming is the suppression or elimination of ICI. In a Single-User (SU) point-to-point scheme, ICI is caused by multi-stream transmission, while in a MU architecture, ICI is the result of the concurrent transmission to multiple users using the same time and frequency resources. The term that is used to characterize this form of processing is transmit precoding or simply precoding. Precoding can be linear [70] or non-linear [71]. This thesis is mainly focused on linear precoding, due to its simplicity and low complexity.

The baseband system equation of a beamforming architecture with linear precoding can be obtained from (2.3), if it is assumed that the transmitted signal vector is linearly processed before its transmission. Thus, the system equation becomes:

$$\mathbf{y} = \mathbf{HPDx} + \mathbf{w}. \tag{2.14}$$

In (2.14), $\mathbf{P}$ is a $N_t \times N_r$ matrix which denotes the linear precoder. In order to constrain the

transmission power to acceptable levels, a diagonal normalization matrix:

$$\mathbf{D} = \mathbf{diag}\left(d_1, \ldots, d_{N_r}\right) \tag{2.15}$$

is employed. Each element d_i, $i = 1, \ldots, N_r$ equals to:

$$d_i = \sqrt{\frac{1}{\|\mathbf{p}_i\|_2^2}}, \tag{2.16}$$

where, $\mathbf{p}_i$ denotes the i-th column of $\mathbf{P}$. In this way, each column of the normalized precoding matrix:

$$\mathbf{P}_{\text{norm}} = \mathbf{PD} \tag{2.17}$$

has unity power. In contrast to (2.3), where $\mathbf{x}$ is a $N_t \times 1$ vector which encodes the binary information, in (2.14), $\mathbf{x}$ is a $N_r \times 1$ vector which has to be linearly processed before its transmission. All the other parameters of (2.14) are the same as (2.3). Note, that the normalization of the linear precoder can be achieved by using alternative approaches, such as the ones in [72, 73]. However, in this thesis, the normalization of the linear precoder is undertaken as described in (2.14)–(2.17). In a MU setup, the previous normalization approach results in a form which is very convenient for the application and analysis of different adaptive power (resource) allocation algorithms.

The most common linear precoding methods, when P-CSIT is available, are ZF and MMSE precoding. In conventional MIMO schemes, ZF (Bezout) precoding is a suboptimal method that offers a good trade-off between complexity and performance [74]. The main characteristic of ZF precoding is the total elimination of the Inter-Channel Interference (ICI) at the receiver, $\mathbf{HP}_{\text{ZF}} = \mathbf{I}_{N_r,N_r}$, where $\mathbf{P}_{\text{ZF}}$ is the ZF precoder and $\mathbf{I}_{N_r,N_r}$ is the $N_r \times N_r$ identity matrix. The ZF precoding matrix is simple the pseudo-inverse of the channel matrix $\mathbf{H}$. Using the Singular Value Decomposition (SVD), the precoding matrix is expressed as:

$$\mathbf{P}_{\text{ZF}} = \mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^H, \tag{2.18}$$

where, $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$ and $(\cdot)^{-1}$ denotes the inverse matrix. Both $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices with size $N_r \times r$ and $N_t \times r$, respectively. Furthermore, $\mathbf{\Sigma}$ is a $r \times r$ diagonal matrix whose main diagonal represents the r real singular values σ_i of $\mathbf{H}$. Here, r denotes the rank of the

channel matrix $\mathbf{H}$. A different way to express the ZF precoder is given below:

$$\mathbf{P}_{\text{ZF}} = \mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H\right)^{-1}. \tag{2.19}$$

An alternative method suitable for rank-deficient channel matrices is MMSE (regularized ZF) precoding. The MMSE precoder is given as:

$$\mathbf{P}_{\text{MMSE}} = \mathbf{H}^H \left[\mathbf{H}\mathbf{H}^H + \mu \mathbf{I}_{N_r,N_r}\right]^{-1}, \tag{2.20}$$

where, μ is the regularization factor. Usually, μ is defined as $\mu = \frac{N_t \sigma_w^2}{P_{\mathbf{s}}} = \frac{N_t}{\text{SNR}}$ [72]. In conventional MIMO, MMSE precoding outperforms ZF precoding in low SNRs and approaches the ZF performance in high SNRs.

When the linear precoder $\mathbf{P}$ is designed based on the ZF or MMSE criterion, the system equation of (2.14) effectively represents: i) a SU point-to-point MIMO system, where all the N_r receive antennas belong to the same receiver; or ii) a MU MIMO system where the N_r receive antennas are divided between multiple distributed receivers. In the first case, N_r parallel non-interfering streams are formed from the transmitter to the single receiver. Whereas, in the second case, the N_r parallel non-interfering streams are allocated between the existing users. Usually, each user is allocated a number of streams equal to its receive antennas. Here, implicitly it is assumed that $N_r = \sum_{i=1}^{N_u} K_i$, where N_u is the number of users and K_i is the number of antennas of user i. The latter case is a form of the well known broadcast channel [75], which is a SDMA technique. In both systems, provided that ICI is eliminated (ZF) or suppressed (MMSE), detection can be done efficiently using the ML principle.

2.2.5 Virtual MIMO Architectures

The previous subsections demonstrate how MIMO systems achieve multiplexing, diversity, SNR gains, and ICI suppression. However, the accommodation of multiple antenna elements in a single transmitter or a single receiver could be challenging due to cost and implementation factors. An efficient approach which overcomes these challenges is cooperative communication. In cooperative communication, multiple remote nodes, termed as RNs, cooperate in order to establish communication from a SN to a DN. The concepts of cooperation and relaying between remote nodes are old and partly rely on ad hoc networks. However, after the rise of MIMO, the use of multiple RNs changed direction toward the formation of VMIMO schemes.

In VMIMO schemes, multiple remote nodes cooperate in order to mimic the operating principle of MIMO. Thus, in addition to the gains obtained from relaying and cooperation, a VMIMO scheme benefits from the virtual multi-antenna transmission.

The main rationale behind VMIMO is to promote low complexity system implementation by combining MIMO and relay techniques. Due to the broad nature of VMIMO, there is a wide range of designs proposed in the literature. From the MIMO point of view, there are designs that achieve diversity gains [76], multiplexing gains [77], beamforming gains [78], or a combination of these. Also, from the relays' point of view, there are several communication protocols. The most common protocols are Amplify-and-Forward (AF) [78], DF [76], Compress-and-Forward (CF) [79], and Estimate-and-Forward (EF) [80]. Furthermore, the previous protocols are applied in two [81] or multiple hops [76].

2.2.6 Hybrid MIMO Architectures

Hybrid MIMO architectures include a large number of systems that can be classified in more than one category from the above. An example of a hybrid MIMO systems is the point-to-point Eigen-BeamForming (EBF) [82]. In EBF, both communication ends deploy linear processing using Channel State Information (CSI). Furthermore, depending on the number of the deployed parallel streams, the system diversity order varies [82]. Thus, EBF can be classified as a SMX, a diversity achieving and a beamforming architecture. In this case, there is a trade-off between different gains. Similarly, a point-to-point linearly precoded MIMO system, such as the one described by (2.14), is categorized as a beamforming spatially multiplexed architecture.

2.2.7 Alternative Categorization of MIMO Architectures

In addition to the categories of multi-antenna communication presented in Sections 2.2.2–2.2.6, MIMO systems can be categorized by the following different approaches. For instance, there are SIMO schemes where the transmitter has one antenna while the receiver has multiple antennas. A MISO system is the reverse of a SIMO system. In MISO, the receiver uses one antenna, whereas the transmitter deploys multiple antennas. Another form of distinction of MIMO systems is between narrowband and broadband schemes. In narrowband MIMO schemes, the wireless channel is frequency flat, while in the broadband MIMO schemes, the wireless channel has a frequency selective nature. The presence of CSI at the transceiver distinguishes MIMO

in open loop, closed loop, and non-coherent techniques. An open loop MIMO scheme utilizes channel knowledge only at the receiver. In contrast, a closed loop MIMO scheme benefits from the presence of CSIT. Architectures which do not employ CSI belong to non-coherent MIMO techniques. Furthermore, MIMO schemes can be divided between co-located and distributed architectures. Co-located MIMO communication includes systems where the antenna array of a communicating end (transmitter or receiver) is placed in a certain point (device). In contrast, in distributed MIMO, the antenna arrays are distributed in a wide geographical area and are connected via backhaul connections. Usually, distributed MIMO architectures aim to obtain macro-diversity gains. Finally, there is a distinction between SU and MU MIMO. The category of SU MIMO includes systems which establish a point-to-point communication between a single transmitter and a single receiver. Whereas, in MU MIMO, multiple users are concurrently served through the same time and radio resources via SDMA, either using the broadcast channel (downlink) or a MAC technique (uplink).

2.2.8 New Trends in MIMO Communication

As MIMO communication develops, its spectral efficiency approaches more challenging limits and new trends in the deployment of MIMO appear. Besides, MIMO communication seems to be an important part of the future Fifth Generation (5G) of cellular communication systems [83]. The last few years, a wide range of concepts that utilize MIMO have been researched, such as Coordinated MultiPoint (CoMP) [84]; massive MIMO [85, 86]; millimeter wave communication [67–69]; hybrid optical and RF systems; energy harvesting [87, 88]; BS densification; and EE [12, 89, 90].

In a cellular system, providing the cell-edge users with sufficient spectral efficiency, without sacrificing valuable radio resources, is a long-standing but still open problem. The spectral efficiency of cell-edge users is limited due the existing interference from neighboring cells. An approach that MIMO communication follows in order to provide a solution to this problem is coordinated SDMA, also termed as CoMP transmission. Taking advantage of the space division between different users and by applying the appropriate signal processing technique, CoMP transmission can be deployed both in the downlink and uplink. In CoMP systems, the neighboring cells, instead of creating interference for each other, they coordinate in the sense of precoding (downlink) or joint MU detection (uplink) in order to form a distributed MIMO architecture with increased spectral efficiency.

Very large scale MIMO or massive MIMO communication is a new research field that has recently emerged [86]. In [85], high data rate performance is demonstrated for a multiuser scenario without BS cooperation. The concept of massive MIMO challenges EE gains by using a very large number of low power transmitting antennas. In addition, the employment of a large number of antennas makes the wireless system more robust to the deactivation of a couple of antenna units, when compared with small scale MIMO. From a mathematical point of view, the analysis of massive MIMO can be undertaken using the asymptotic random matrix theory [86]. Thus, the analysis of massive MIMO is simplified compared to small scale systems. In massive MIMO systems, the channel matrix is either very fat or very tall. Hence, usually the channel matrix tends to have a good condition number. Another characteristic of the massive MIMO channel is its extreme size. Thus, the feedback of CSI from the receiver becomes impractical. Therefore, TDD could be a plausible solution. This happens because, in this case, the reciprocity principle becomes valid [86]. In such a scenario, CSIT is acquired via a pilot sequence transmitted from the receiver. Hence, the assumption of P-CSIT approaches to reality.

As the limit of spectral efficiency has to be extended constantly, an efficient and well known approach is to exploit new spectrum bands. This approach is a direct result of the analysis of (2.1), presented in Section 2.1. In agreement with this approach, the exploitation of millimeter wave frequency bands appears to be a promising solution [67]. Further justifications for the deployment of millimeter wave communication are: i) the low cost and low complexity Complementary Metal Oxide Semiconductor (CMOS) technology that can be used; ii) the simple beam-steering antennas that can be deployed at both communicating ends. In addition, the broad available bandwidth can be translated in lower latency, which is of high importance in the design of future 5G [83]. However, due to the high path loss of millimeter wave frequency bands, several challenges need to be addressed. For example, accurate beamforming requires CSI of high quality at both communicating ends.

In the direction of increasing the available bandwidth, the new and innovative frameworks of Visible Light Communication (VLC) [91–94] and Free Space Optics (FSO) communication [95, 96] have been developed. In a wide range of application scenarios of optical wireless communication, objectives such as available bandwidth, spectral efficiency, EE, secrecy, and coverage could have a simple and effective solution due the nature of the communication medium; but, which is the role of MIMO in optical wireless communication? The combined use of optical wireless communication and RF MIMO is able to offer a significant increase in data rate.

Hybrid optical/RF communication is a very promising field for future research. Furthermore, concepts from RF MIMO communication can be adapted in a pure optical framework in order to offer many of the fundamental advantages of MIMO, such as diversity, SMX, and SDMA [64, 97, 98].

As the need for higher spectral efficiency becomes more exacting, network deployment and planning also become vital. In addition to the deployment of MIMO transmission, the network has to become more dense by deploying more BSs. In this way, multi-tier networks are formed. The concepts of densification and offloading are simple and demonstrate clear advantages. As the cell size decreases, the frequency spectrum is better utilized, while fewer users compete for the same resources. However, the parameter interference becomes dominant. In research, mathematical tools from stochastic geometry have been employed in order to study the effect of interference in a random dense network [99, 100]. Furthermore, equally important, parameters are the network cost and energy consumption. It is expected 5G to offer a tremendous performance increase compared with 4G [83]. Concurrently, the cost and energy consumption of the future networks are expected to be lower. Therefore, the use of MIMO communication in such networks needs to be researched. In this context, the deployment of multiple antenna elements promises to offer energy harvesting in a more efficient way [87, 88].

2.2.9 Challenges and Drawbacks of MIMO Architectures

The previous sections demonstrate the advantages of MIMO communication. However, these advantages also incur disadvantages. As the system size scales, a number of challenges need to be addressed, especially, as research moves toward massive MIMO. For example, the signal processing algorithms, which have to be executed in real time, could be a significant challenge. Another challenge is the need for tight synchronization among antenna elements.

Furthermore, multi-antenna elements face a major disadvantage of requiring multiple RF chains. This is a major disadvantage because RF chains are expensive circuits that do not follow Moore's law [11]. Thus, the real system implementation becomes expensive and often impractical. In addition, RF chains are electronic circuits that may reduce the EE of the wireless system. Indeed, RF chains include PAs which are responsible for 50-80% of the total power consumption in the transmitter [18]. The authors of [89, 101] shown that even though conventional MIMO systems offer high data rates, their EE diminishes significantly due to the use of multiple RF chains. Therefore, there is a significant need for MIMO schemes where the

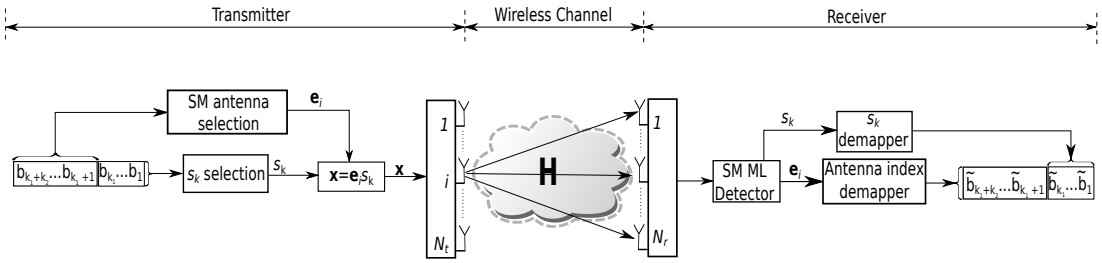


Figure 2.3: *System Model of SM.*

number of RF chains does not scale with the number of deployed antennas. Such an example is envisioned by the authors of [102]. In that paper, a massive MIMO system with few high power PAs and multiple low power PAs, plugged in each antenna, is proposed. An alternative, novel, and very promising MIMO scheme that deploys multiple transmit antennas and requires a single RF chain in the transmitter is SM [14–16, 103]. The rest of this thesis focuses on the evaluation and extension of the concept of SM as a low complexity, energy efficient MIMO scheme.

2.3 Spatial Modulation: Rethinking the Space Resources

Unlike conventional MIMO schemes, SM exploits multiple transmit antennas in a way that only a single RF chain is required. Fig. 2.3 shows the underlying concept of SM. During the signaling period, the bits to be transmitted are divided into two blocks. The first block is encoded using a conventional signal constellation diagram such as QAM. Provided that the number of antennas is a power two and each antenna is allocated a unique binary index, the second block of bits is used for the selection of the single transmitting antenna. The transmitting antenna possesses the binary index which corresponds to the second block of bits. Well known advantages of SM are the avoidance of Inter-Antenna Synchronization (IAS) at the transmitter and ICI at the receiver [16]. In addition, the detection at the receiver is performed using a low-complexity (single stream) ML detector [104], which jointly detects the conventional constellation point and the index of the transmitting antenna. Note, that even though SM employs a single stream detector, it is able to achieve a multiplexing gain.

2.3.1 Operating Principle of Spatial Modulation

This section presents in detail the operating principle of SM. A point-to-point MIMO system with N_t transmit antennas and N_r receive antennas is considered. It is further assumed that only the receiver possesses P-CSIR. When the wireless channel is flat, the general MIMO system equation is given in (2.3).

The basic concept of SM is the transmission of a standard symbol form the single activated transmit antenna. Thus, if the number of transmit antennas is configured to be a power of two, the transmitted signal vector is formulated as $\mathbf{x} = \mathbf{e}_i s_k$, where $\mathbf{e}_i = [0, ..., 0, 1, 0, ..., 0]^T$, $i = 1, ..., N_t$, is a $N_t \times 1$ vector whose elements are zero except the i-th element which has the value of 1. Clearly, the position of the non-zero element of $\mathbf{e}_i$ corresponds to the position of the activated antenna. In addition, s_k is a standard symbol drawn from a M-ary constellation such as QAM or PSK. In this case, the baseband system equation of SM becomes:

$$\mathbf{y} = \mathbf{H}\mathbf{e}_i s_k + \mathbf{w}, \tag{2.21}$$

where, $\mathbf{y}$, $\mathbf{H}$, and $\mathbf{w}$ are defined just as in (2.3). Therefore, a total of $k_{\text{SM}} = k_1 + k_2$ bits per channel use (bpcu) are conveyed from the transmitter to the receiver. The first $k_1 = \log_2 N_t$ bits are used for the selection of the single activated antenna, and the other $k_2 = \log_2 M$ bits are used for the selection of the transmitted symbol s_k.

Provided that each symbol period SM requires the activation of a single transmit antenna, a real system implementation could be undertaken using: i) a single RF chain which creates the waveform to be transmitted and corresponds to s_k; and ii) a quick antenna switch which selects the activated antenna. However, is it possible to implement an antenna switch which operates in the symbol rate? The authors of [69] report that, even for a millimeter wave communication scenario, it is possible to implement such a switch based on several solid-state technologies.

At the receiver side, the objective is to detect the transmitted constellation point and the index of transmitting antenna. In published research, several sub-optimal approaches can be found [103, 105–107]. However, the optimal ML detector of SM is given in [104]. This detector is formulated as:

$$(\hat{i}, \hat{s}_k) = \arg \min_{i, s_k} \|\mathbf{y} - \mathbf{H}\mathbf{e}_i s_k\|_2^2. \tag{2.22}$$

The index of the activated antenna and the transmitted standard symbol s_k are jointly detected.

Although, SM is able to achieve a multiplexing gain, (2.22) shows that the optimum ML detector of SM is a single stream.

2.3.2 Existing Point-to-Point Architectures Based on Spatial Modulation

The concept of space modulation dates from the early 2000s, when it was independently proposed by different researchers [108–110]. However, it gained popularity after the publication of [14, 103]. Today, most of the existing space modulated architectures are inspired by or are variants of [14, 103, 104, 111, 112]. The main objective of these papers is to promote a trade-off between spectral efficiency and low complexity. Complexity here denotes: i) the encoding/decoding complexity; and ii) number of deployed RF chains in the transmitter.

In [103], the concept of SM is presented using a low complexity suboptimal detector, while [104] gives the optimum ML detector of SM and its performance analysis. The extension of SM in an Orthogonal Frequency-Division Multiplexing (OFDM) system is conducted by [14].

A low complexity and low rate variant of SM, termed Space Shift Keying (SSK), is proposed in [111]. The main difference between SSK and SM is that it does not use a conventional constellation diagram and conveys information only via the index of the activated antenna. In this way, the bit-stream to be transmitted during a symbol period is encoded in the index of a single transmitting antenna. At the receiver side, the transmitted bitstream is reconstructed through the detection of the index of the transmitting antenna. The system equation of SSK is given from (2.21) by setting $s_k = 1$. Thus, taking advantage of the sparsity of $\mathbf{e}_i$, (2.21) is transformed to:

$$\mathbf{y} = \mathbf{h}_i + \mathbf{w}, \tag{2.23}$$

where, $\mathbf{h}_i$ is the i-th column of $\mathbf{H} = [\mathbf{h}_1, \ldots, \mathbf{h}_{N_t}]$. Similarly, the ML detector of SSK is derived from (2.22) by setting $s_k = 1$ and searching only through the indices i:

$$(\hat{i}) = \arg\min_{i} \|\mathbf{y} - \mathbf{h}_i\|_2^2. \tag{2.24}$$

In this case, SSK is capable of transmitting $k_{\text{SSK}} = \log_2(N_t)$ bpcu.

As indicated before, a mandatory requirement for pure SM and SSK is the number of transmit antennas to be a power of two. However, this requirement may be too restrictive in practical scenarios. The first scheme to overcome this restriction is Generalized Space Shift Keying (GSSK)

[112]. In GSSK, every symbol period, a number of transmit antennas are activated in order to encode the transmitted bit-stream. The spectral efficiency of GSSK is $k_{\text{GSSK}} = \lfloor \log_2 (N_c) \rfloor$ bpcu, where N_c is the number of possible combinations of active antennas. Provided that, during a symbol period, the number of active antennas N_a is constant, it holds that $N_c = \binom{N_t}{N_a}$. Following a similar approach, the authors of [113] extend the concept of GSSK by encoding additional information in the transmission of a conventional constellation point. This scheme is known as Generalized Spatial Modulation (GeSM). A different methodology that eliminates the need for the number of antennas to be a power of two is applied in [114, 115]. In these papers, fractional bit rates are achieved using the modulo conversion.

The concepts of SM and SMX are combined in [19, 116]. In this way, a generalized multi-stream space modulation scheme is formed. In [19, 116], every symbol period, a subset of the available transmit antennas is activated, with each antenna transmitting a different conventional symbol. At the receiver, detection can be undertaken using the ML principle [116]. Alternatively, as shown by the authors of [19], detection can be undertaken using a V-BLAST-like suboptimal detector. Due to its operating principle, the scheme of [19, 116] increases the detection complexity and the number of RF chains employed at the transmitter.

As with every other MIMO scheme, SM has a performance degradation in the presence of channel correlation. In order to mitigate this limitation, the authors of [117] propose a coding scheme for SM based on Trellis Coded Modulation (TCM). Unfortunately, the coding method of [117] does not provide any performance gain compared to the uncoded SM when the wireless channel is uncorrelated. However, [118] presents a form of TCM which allows coded SM to have performance improvements both in uncorrelated and correlated channels.

The performance analysis of conventional SM shows that it is not able to achieve transmit diversity [119, 120]. Due to this conclusion, there has been several approaches that modify the transmission mechanism of SM in order to achieve transmit diversity gains [121–124]. The first incorporation of STBC in a spatially modulated system is conducted in [121] by using the Alamouti code. In [121], it is shown that the spatially modulated Alamouti codeword has a normalized rate higher than one, without increasing significantly the system complexity. In addition, simulation results demonstrate that the spatially modulated Alamouti code has a better performance than the State-of-the-Art (SotA). An approach for constructing orthogonal spatially modulated codes is introduced in [124]. An alternative methodology is followed in [122, 123], where the concept of time orthogonal shaping filters is combined with SM and

SSK, respectively, in order to offer transmit diversity. Both papers demonstrate that space modulation schemes with time orthogonal shaping filters are able to offer transmit diversity gains with higher rate and lower complexity than the SotA. Other architectures that promote transmit diversity for SM include those presented in [125–130].

The deployment of CSIT is shown to offer significant performance enhancements in space modulated systems. The optimal power allocation for a SSK system with two transmit antennas is conducted in [131]. A form of constellation design for SSK and GSSK, under the presence of CSIT, is presented by [132]. Furthermore, several papers deal with the problem of antenna selection for SM [125–128]. The concept of link adaption for SM, based on CSIT, is studied in [133–136]. A closed-loop scheme which applies the concept of SM at the receiver side is proposed in [20] and investigated under I-CSIT in [137].

Another broad category of wireless MIMO systems inspired by SM and based on the dispersion matrix encoding designs is proposed in [130, 138]. In these papers, information is encoded via the deployment of space-time domain matrices which spread the transmit bits across space and time. In this way, the concept of SSK is extended to the space-time domain. Also, in [130, 138], it is shown that the proposed scheme is able to adjust its operating mode in order to facilitate single or multiple transmission streams. In this way, a trade-off between rate and complexity is established. Furthermore, the work of [138] demonstrates that a transmit diversity gain can be obtained if the appropriate dispersion matrix is selected. The design of spreading matrices for the schemes of [130, 138] is extensively studied in [139–141].

2.3.3 Existing Multiuser Space Modulated Architectures

Section 2.3.2 demonstrates the wide range of existing spatially modulated architectures for point-to-point communication. However, is it possible for SM to be incorporated in a MU system? Just like any other MIMO physical layer technique, SM and its variants can be combined with a multiple access scheme such as TDMA, FDMA, or OFDMA in order to form a MU system.

A new trend in MIMO communication promotes systems where multiple users are aggressively allocated in the same time and frequency resources. Usually, this is accomplished via SDMA techniques. Alternatively, multiple users are accommodated without addressing the existing interference or via the use of interference aware detectors. In the first case, the system design

establishes a trade-off between performance and aggregate rate, while in second case, the trade-off is between complexity and system performance.

In this context, the authors of [142] study the performance of SSK in a multiuser scenario both for the interference aware and unaware detectors. Similarly, the authors of [143, 144] study the performance of MU SM in the uplink when the detector at the receiver is interference aware or unaware. In general, the interference unaware detector faces a BER error floor. However, the interference aware detector does not result in an error floor, but it has a performance degradation compared to the SU case.

The design of a SDMA-based spatially modulated scheme is a challenging task, especially for the downlink where the broadcast channel has to be established. In the downlink, the main difficulty of the design of such a system originates from the operating mechanism of SM. In more detail, due to the activation of a single transmit antenna (or a portion of the available transmit antennas) and the way that information is conveyed, the deployment of interference reduction, elimination, or manipulation techniques becomes difficult. However, the authors of [24, 25] managed to incorporate SM in the broadcast channel. In these papers, every user is allocated a portion of the transmit antennas. Non-interfering SM-based communication is established via the use of a linear precoding matrix based on the ZF principle. The scheme of [24, 25] requires both CSIT and CSIR.

2.3.4 Existing Virtual MIMO Architectures Based on Spatial Modulation

Over the recent years, there have been several schemes that extend the concept of SM/SSK to a VMIMO distributed scenario, both for the AF and DF protocols. The authors of [26] study the performance analysis of a SSK-based AF relay network. In [145], using the concept of SSK, a cooperative transmission scheme is considered, both for the AF and DF protocols. The generalization of SSK in a two-way AF relays network is conducted in [146]. The adoption of SSK in a DF cooperative system is considered in [147]. Furthermore, the extension of SSK to a multi-branch and multi-hop scenario is studied in [148]. Also, the extension of Space Time Shift Keying (STSK) [129], which is a novel SM-like scheme, in a distributed VMIMO scenario is conducted in [149].

In [150], multi-antenna elements are deployed in such a way that form a dual hop scheme based on SM. The application of distributed SM to the uplink is studied in [151]. Also, [152]

proposes a SSK scheme again for the uplink. In [153], using the ergodic capacity and the outage probability, it is shown that space modulation VMIMO schemes are able to achieve high throughput. Another extension of SM in a low complexity cooperative DF VMIMO architecture is presented in [27]. The integration of STBC in a distributed SM system is reported in [154].

Unfortunately, all distributed and uncoded space modulation VMIMO schemes that employ multiple uncoordinated RNs suffer from the erroneous activation of multiple RNs during the relaying phase. Thus, a BER performance degradation occurs. In order to overcome this problem, three strategies are followed in the literature: i) coordination between the RNs in order to ensure the activation of a single RN [155] (which is similar to the scenario of a single RN with multiple antennas [150]); ii) block-based transmission which allows the use of error correction codes at the RNs [145, 148]; and iii) advanced and error-aware detection techniques at the DN, which usually results in high complexity [151, 154].

2.3.5 Motivation of This Thesis

Research on MIMO communication systems has shown that the use of multi-antenna elements offers a higher data rate and lower BER without additional RF resources. Unfortunately, in most MIMO architectures, the system complexity increases as the number of antenna elements scales. Thus, the real system implementation becomes challenging and the energy consumption is increased. Furthermore, in MIMO architectures, the computational complexity becomes a major issue, at both communication ends. Therefore, the question whether it is possible to design a multi-antenna communication scheme which benefits from the advantages of multiple antennas while retains low complexity, becomes vital. Clearly, the same question extends for all multi-antenna deployments, such as VMIMO and MU communication.

SM is a MIMO scheme that exploits multiple antennas in a different way than conventional MIMO and aims to promote low complexity and EE. Due to its transmission mechanism, SM incorporates a single RF chain configuration at the transmitter, while its receiver is able to employ a single stream (low complexity) optimum ML detector. SM is an established MIMO scheme. In more detail, the deployment of SM and its variants in an open loop point-to-point scenario have been extensively researched [119, 156–159]. Also, although the incorporation SM in a MU scenario already exists [24, 25, 142–144], it is not yet considered mature, especially in the downlink. In addition, as demonstrated in Section 2.3.4, several approaches extend the concept of SM in a distributed cooperative VMIMO framework [26, 27, 129, 145–154].

The effect of increasing the number of transmit antennas, and consequently the number of PAs, to the power consumption of the transmitter is studied in [12, 13, 18, 28]. In more detail, the authors of [12, 13] reveal that the main energy consumers in a cellular network are the deployed BSs. Furthermore, [12, 13, 18, 28] demonstrate that the major consuming units of a BS are the deployed PAs. Therefore, as the number of PAs increases, the energy consumption of a BS also increases. Given that the transmission mechanism of SM requires a single RF chain (in single carrier transmission), it can be concluded that a BS which employs SM consumes less energy than a BS which employs conventional MIMO. Note that a similar analysis can be undertaken also for a mobile terminal employing SM. However, this conclusion does not imply directly that SM is more energy efficient than conventional MIMO. Thus, the aim in Chapter 3 is to quantify the EE, in terms of Mbits/J, of a BS deploying SM and its power supply, in W, using realistic system assumptions. Also, in the same chapter, a comparison between SM and conventional MIMO, in terms of EE (Mbits/J) and total power supply for the same data rate, is provided.

Similar to the concept of conventional SM, the author of [20] proposes a point-to-point closed-loop MIMO scheme which applies the principle of SM at the receiver side. In this scheme, using MIMO linear precoding with CSIT, the transmitter is able to aim to a single receive antenna out of N_r. Given that the number of receive antennas N_r is a power of two and each receive antenna is allocated a unique binary index, the reciprocal of SM, called R-SM, is formed. Similar to conventional SM, in R-SM, a portion of the transmitted bit-stream is encoded by selecting the appropriate single receiving antenna (the one which is allocated with the corresponding binary index). The rest of the bit-stream is conveyed via a conventional constellation point such as QAM. The extension of R-SM to a scheme which spatially modulates multiple parallel symbol streams to the indices of multiple receiving antennas is conducted in [21]. In this thesis, the term of MSR-SM is used for this scheme. A complete description of R-SM and MSR-SM is given in the corresponding papers and in Chapter 4. Motivated by R-SM and MSR-SM, the objectives of Chapters 4 are to: i) study theoretically the performance of these schemes when ZF precoding with P-CSIT is employed; ii) study their performance assuming realistic scenarios of I-CSIT; iii) propose linear precoding methods which employ I-CSIT in a statistical and worst case form. Finally, in Chapter 5, the aim is to propose a multiuser SDMA scheme based on R-SM or MSR-SM and study its performance.

As described in Section 2.3.4, motivated by the potential of point-to-point SM, research has extended the concept of SM in a distributed framework by incorporating relaying techniques.

The main objective of this research is to form VMIMO architectures, which mimic the operating concept of conventional point-to-point SM. In VMIMO architectures, a SN communicates with a DN via the cooperation of multiple RNs. As a consequence, the main aim of any SM-based VMIMO architecture is to inherit the advantages of conventional SM. Unfortunately, the formation of a distributed SM transmission mechanism is a challenging task when the cooperating nodes are uncoordinated. In such a scenario, due to the independent operation of the cooperating RNs, the retransmitted signal might not belong to the aimed transmission alphabet. For example, in a SM-based scheme, two RNs could retransmit different constellation points. Note that this is not limited only to the SM-based VMIMO architectures. Such an example is [160] which proposes a form of distributed Alamouti STBC. In such a scheme, the orthogonality of the formed distributed Alamouti codeword is not guaranteed when the deployed nodes are uncoordinated. In published research, three different strategies are followed in order to solve this drawback. The first strategy enforces communication between the cooperating RNs. Therefore, additional resources are required. An alternative strategy is the adoption of block based transmission. In this way, error correction coding is utilized in order to ensure that the relayed signal belongs to the employed legal alphabet. Therefore, the complexity of the RNs is increased. The last strategy requires the use of advanced error aware detection techniques at the DN. Thus, the complexity of the DN is increased. Motivated by this framework, this thesis proposes and studies the performance of a VMIMO architecture based on SM which is suitable for the downlink. As shown in Chapter 6, the new architecture resolves the native problem of retransmission of illegal signals from the RNs by deploying linear precoding at the SN. In this way, the complexity of the RNs and the DN remains unaffected. In fact, this is desirable in the downlink, where the SN is a BS with less strict complexity constraints, while the RNs and DN are remote terminals with strict complexity constraints.

2.4 Summary

In this chapter, a brief historical review of modern wireless communication was presented. It demonstrated the way that the scarcity of the available RF bandwidth shaped the evolution of modern wireless systems. Also, in Chapter 2, it was shown why even though the concept of MIMO communication exists since early 1970s, it was only after its information theoretical analysis, in late 1990s, that it gained its popularity among the researchers. Also, the most important categorization of MIMO systems was provided. The main rationale behind each

category was described. In addition, the drawbacks and challenges of MIMO communication systems and their new trends were reviewed in Chapter 2. In this context, the concept of SM was introduced. The existing variants of SM and their main advantages and limitations were presented. Finally, the main motivation of this thesis was provided.

Chapter 3
Energy Evaluation of Spatial Modulation at the Downlink

3.1 Introduction

The aim of this chapter is to study the EE of SM at different BSs taking into account the total power consumption. Compared to conventional MIMO schemes, SM benefits from a single RF chain configuration which results in decreased power supply (W), higher EE (Mbits/J), and reduced complexity. In Chapter 3, using the fundamental limits of channel capacity, it is shown that SM achieves a range of average data rates with only a fraction, which can be as low as 24% for four transmit antennas, of the total power supply of conventional MIMO. In addition, it is demonstrated that the EE of the studied schemes is maximized for a certain average data rate and that SM achieves the highest EE among them. Finally, it is noted that a BS employing SM can be up to 67% more energy efficient compared to a BS under a conventional MIMO transmission scheme, for four transmit antennas.

3.2 Information-Theoretic Analysis of Spatial Modulation

Since SM is a relatively new MIMO scheme, its information-theoretic analysis has attracted significant attention [130, 161–163]. The first evaluation of the channel capacity of SM is provided by the authors of [161]. In that paper, under the assumption that the transmitted signal is selected as an independent and identically distributed (i.i.d.) complex Gaussian RV, the SM rate is expressed as the sum of two parts [161]:

$$C_{\text{SM}} = W\left(C_1 + C_2\right). \tag{3.1}$$

Here, it is reminded that W is the available bandwidth. The main rationale for expressing the rate of SM as (3.1) is its transmission mechanism. SM conveys information via the transmitted conventional signal and the index of the activated antenna. Thus, in (3.1), C_1 denotes the

capacity that is linked to the conventional signal in the complex signal space and partially to the spatial symbol (antenna index), and is equal to:

$$C_1 = \frac{1}{N_t} \sum_{i=1}^{N_t} \log_2 \left(1 + \frac{P \|\mathbf{h}_i\|_2^2}{\sigma_w^2} \right). \tag{3.2}$$

Note that the derivation of C_1 is based on the analysis presented in [6]. It is not difficult to see that the capacity that is linked with the conventional signal, C_1, reduces to the capacity of the so-called "spatial cycling using one transmit antenna at a time" scheme of [6, Section 4.E]. Additionally, C_2 in (3.1) exclusively describes the additional rate achieved by SM by using the index of the single activated antenna for transferring information.

The direct evaluation of C_2 is a difficult task [89, 101, 162]. For this reason, the authors of [162] provide the following lower and upper bounds of C_2:

$$\log_2 (N_t) - \frac{1}{N_t} \sum_{i=1}^{N_t} \log_2 \left(\sum_{j=1}^{N_t} \frac{e \left[\|\mathbf{h}_i\|_2^2 \frac{P}{\sigma_w^2} + 1 \right]}{\left[\|\mathbf{h}_i\|_2^2 + \|\mathbf{h}_j\|_2^2 \right] \frac{P}{\sigma_w^2} + 2} \right) \leq C_2 < \log_2 (N_t). \tag{3.3}$$

In [161, 162], the capacity of SM is derived in the Shannon sense (channel capacity). In order to gain a better understanding, the authors of [130] provide the capacity of a wide range of space modulated schemes in the Discrete input Continues output Memoryless Channel (DCMC). Finally, the authors of [163] derive the mutual information of SM for a $N_t \times 1$ configuration, when the transmitted signal belongs to a finite alphabet.

In the rest of this chapter, the channel capacity of SM is lower bounded as:

$$C_{\text{SM}} > W C_1. \tag{3.4}$$

This lower bound is valid because the presented results serve as a lower bound on the data rate and EE performance.

3.3 Power Consumption of a Base Station Employing Spatial Modulation

A modern multi-antenna BS consists of multiple transceivers. Each transceiver contains: i) an Antenna Interface (AI); ii) a PA; iii) a Baseband Interface (BI); iv) a Direct Current (DC)-DC power supply; v) an AC-DC supply; and vi) a cooling system. Higher traffic load increases the power consumption on some of these components, whereas others keep their power consumption constant. The total power consumption of this type of BS is load dependent mainly because of the PAs [17].

In the following, the major components of a transceiver and their effect to the power consumption are analyzed. The power consumption of the AI is characterized by power losses which occur due to the feeder, antenna bandpass filters, duplexers, and matching components.

The main power consumer in a transmitting BS is the deployed PAs. In particular, PAs are responsible for the 50-80% of the total power consumption in a BS [18]. The most energy efficient transmission point of a PA is its maximum transmission power. However, due to non-linearities and the transmission of a signal with a time varying envelope, usually, the PA operates close to its linear region. Thus, its efficiency, η_{PA}, is reduced resulting in an increased power consumption. In [17], the power consumption of a PA is expressed as :

$$P_{\mathrm{PA}} = \frac{P_{\mathrm{t}}}{\eta_{\mathrm{PA}}(1-\sigma_{\mathrm{feed}})}, \tag{3.5}$$

where, P_{t} is the RF transmit power per antenna and σ_{feed} is the feeder's loss.

As stated in [28], the BI is responsible for performing the digital up-conversion and down-conversion, modulation and demodulation, digital pre-distortion in the downlink, signal detection in uplink, equalization, channel coding and channel decoding. In general, the power consumption of the BI corresponds to the 5-15% of the whole BS power supply [18]. Furthermore, additional power losses of the order of 5-10% of the whole power supply take place at the DC-DC and AC-DC power supply [18]. Finally, the 50-80% of the whole power supply is consumed for cooling [18].

Given that the power supply of a BS scales as a function of the number of the deployed transceivers, the authors of [17] relate the RF power transmission of a BS to its power sup-

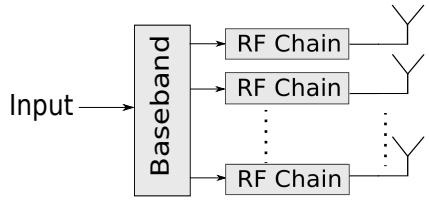


Figure 3.1: *Basic block diagram of a conventional MIMO transmitter.*

ply as:

$$P_{\text{supply}} = N_{\text{RF}} \frac{P_{\text{PA}} + P_{\text{RF}} + P_{\text{BI}}}{(1-\sigma_{\text{DC}})(1-\sigma_{\text{MS}})(1-\sigma_{\text{cool}})}. \tag{3.6}$$

In (3.6), P_{RF} is the power consumption of the Small-Signal RF Transceiver (RF-TRX) which is composed from a receiver for the uplink and a transmitter for the downlink. Also, P_{BI} stands for the power consumption of the BI. Finally, σ_{DC}, σ_{MS}, and σ_{cool} are the loss factors of the DC-DC and AC-DC power supply, and cooling, respectively. Typical values for the previous parameters are given in [17].

The EARTH power model is a very simple and elegant model that simplifies (3.6) and relates the transmitted power of a BS to the total power consumed [28]. Taking into account that some elements of a BS have a traffic load depended behavior and some others have a constant power consumption, the EARTH power model expresses the total power consumption of a BS as an affine function of the RF transmit power:

$$P_{\text{supply}} = \begin{cases} N_{\text{RF}} P_0 + m N_{\text{RF}} P_{\text{t}}, & 0 < P_{\text{t}} \leq P_{\text{max}} \\ P_{\text{sleep}}, & P_{\text{t}} = 0. \end{cases} \tag{3.7}$$

In (3.7), P_{supply} denotes the total power supplied to the BS; N_{RF} denotes the number of RF chains at the BS; P_0 is the minimum power consumption per RF chain when the BS is active; m represents the slope of the load dependent power consumption and can be expressed as, $m = \frac{1}{\eta_{\text{PA}}(1-\sigma_{\text{feed}})(1-\sigma_{\text{DC}})(1-\sigma_{\text{MS}})(1-\sigma_{\text{cool}})}$, by relating (3.7) with (3.6); P_{t} is the RF transmit power per antenna; and P_{max} is the maximum transmitted power of an antenna. Finally, P_{sleep} is the power consumption when the sleep mode is applied. The introduction of the sleep mode is able to offer significant energy consumption benefits by deactivating a set of elements of a BS [164]. Note that the second part of (3.7), $mN_{\text{RF}}P_{\text{t}}$, corresponds to the power consumption for the total RF power transmission.

Considering (3.7), it can be seen that the number of RF chains N_{RF} affects P_{supply} in two ways. The employment of more RF chains results in: i) the increase of the load independent power

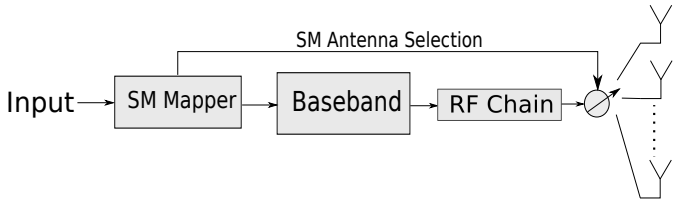


Figure 3.2: *Basic block diagram of a SM transmitter.*

BS type	P_0 (W)	m	P_{max} (W)	P_{sleep} (W)
Macro	118.7	2.66	40.0	63
Micro	53.0	3.1	6.3	-
Pico	6.8	4.0	0.13	-
Femtocell	4.8	7.5	0.05	-

Table 3.1: *Power Model parameters for different BS (SotA 2010).*

consumption of $N_{\text{RF}}P_0$; and ii) the increase of the RF transmit power with a consequential increase of P_{supply}. Note that, in the published research, a number of alternative power consumption models for a BS exist, such as the ones in [165, 166]. However, in this thesis, the focus is on the EARTH power model due to its completeness and elegance. In addition, the EARTH power model parameters are available for different types of BSs in [28].

Unfortunately, as shown in Fig. 3.1, conventional MIMO architectures require multiple RF chains at the transmitter. In contrast, SM is a single RF chain scheme (Fig. 3.2), and although it employs multiple antennas at the transmitter, the number of the utilized RF chains is $N_{\text{RF}} = 1$. In order to quantify the power consumption gain of a multi-antenna BS employing SM ($N_{\text{RF}} = 1$) compared to a BS employing conventional MIMO ($N_{\text{RF}} = N_t$), the EARTH power model of (3.7) is employed. For the case of the same total transmitted power between SM and conventional MIMO, using (3.7), it can be shown after some arithmetic manipulations that:

$$P_{\text{supply}}^{\text{SM}} = P_{\text{supply}}^{\text{MIMO}} - (N_{\text{RF}} - 1)\, P_0. \tag{3.8}$$

Whereas, for the case where the power transmitted from every antenna in conventional MIMO is equal to the power transmitted in SM, it can be shown that:

$$P_{\text{supply}}^{\text{SM}} = \frac{P_{\text{supply}}^{\text{MIMO}}}{N_{\text{RF}}}. \tag{3.9}$$

In (3.8) and (3.9), $P_{\text{supply}}^{\text{SM}}$ denotes the total power supply of SM and $P_{\text{supply}}^{\text{MIMO}}$ denotes the total power supply of a conventional MIMO scheme.

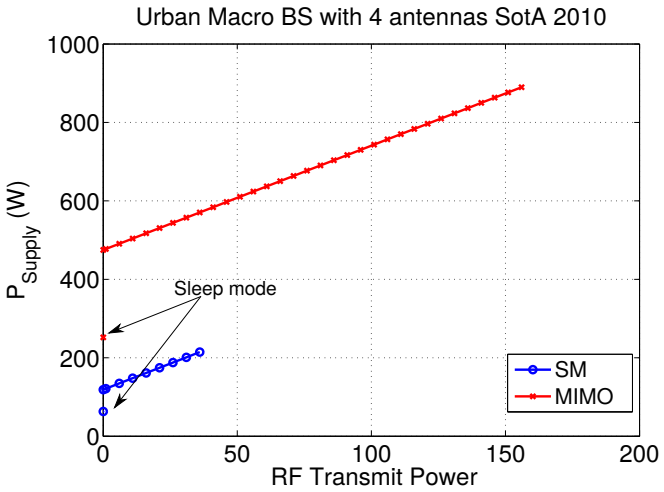


Figure 3.3: *RF transmit power versus P_{supply} for a macro BS with $N_t = 4$ transmit antennas (SotA 2010).*

In order to visualize the power saving of a BS employing SM compared to conventional MIMO, for the same transmit power and using the model parameters from Table 3.1, Fig. 3.3 depicts the RF transmit power versus P_{supply} for a macro BS with $N_t = 4$ transmit antennas. The model parameters of Table 3.1 represent the EARTH model for the year of 2010 [28]. In [28, 90], there is a future prediction for the EARTH power model in year 2014. The main aim of this chapter is to present the EE of SM without taking into account any future hardware improvement. Therefore, in the next sections, Table 3.1 is used and is referred as State-of-the-Art 2010 (SotA 2010) [28].

3.4 Rate and Energy Efficiency Results

In this Section, using the EARTH power model, Monte Carlo simulations are deployed in order to evaluate the data rate and EE of different BSs (macro, micro, pico, and femtocell) in the downlink of a single cell. The studied schemes are: i) SM; ii) STBC [29]; iii) MISO with a single transmitting stream [6]; and iv) MIMO with N_t transmitting streams and channel knowledge only at the receiver [30]. Note that the multiple access scheme is TDMA. Because in this chapter the focus is on the limits, the approximation of the SM channel capacity is deployed in order to obtain the average data rate and EE. For the case of SM, the approximation of the capacity is given in (3.4), whereas for the benchmark techniques the capacity is given below.

Thus, provided that SM requires a single RF chain, the incorporation of (3.7) in (3.4) gives:

$$C_{\text{SM}}(P_{\text{supply}}) = WC_1(P_{\text{supply}}), \tag{3.10}$$

where,

$$C_1(P_{\text{supply}}) = \frac{1}{N_t}\sum_{i=1}^{N_t}\log_2\left(1+\frac{(P_{\text{supply}}-P_0)\|\mathbf{h}_i\|_2^2}{m\sigma_w^2}\right). \tag{3.11}$$

For the case of STBC, the capacity is written as:

$$C_{\text{STBC}} = W\left[R_{\text{STBC}}\ \log_2\left(1+\frac{P}{\sigma_w^2 N_t}\|\mathbf{H}\|_F^2\right)\right], \tag{3.12}$$

where R_{STBC} is the rate of a certain STBC code [167] and $\|\cdot\|_F^2$ is the Frobenius. For complex transmit signals, the rate of STBC is $R = 1$ only for the special case of $N_t = 2$. Generally, $R_{\text{STBC}} = \frac{1}{2}$ for $N_t > 2$ with an exception of three and and four transmit antennas, where $R_{\text{STBC}} = \frac{3}{4}$ [29]. Hence, provided that a BS employs STBC and $N_{\text{RF}} = N_t$, the capacity with respect to the EARTH power model is given as:

$$C_{\text{STBC}}\left(P_{\text{supply}}\right) = W\left[R_{\text{STBC}}\ \log_2\left(1+\frac{P_{\text{supply}}-N_tP_0}{m\sigma_w^2 N_t}\|\mathbf{H}\|_F^2\right)\right]. \tag{3.13}$$

Furthermore, the capacity of a MIMO system with N_t transmitting streams and CSI only at the receiver is given in (2.4). An upper bound of (2.4) is given in [30, eq. 8.17] as:

$$C_{\text{MIMO}} \leq W\left[\min\left(N_t,N_r\right)\log_2\left(1+\frac{P}{\min\left(N_t,N_r\right)\sigma_w^2 N_t}\|\mathbf{H}\|_F^2\right)\right], \tag{3.14}$$

where, $\|\mathbf{H}\|_F^2 = \sum_{i=1}^{\min(N_t,N_r)}\lambda_i^2$. Note that λ_i are the singular values of $\mathbf{H}$. If (3.7) is combined with (3.14), the capacity that can be achieved with respect to the BS power supply is bounded as:

$$C_{\text{MIMO}}\left(P_{\text{supply}}\right) \leq W\left[\min\left(N_t,N_r\right)\log_2\left(1+\frac{P_{\text{supply}}-N_tP_0}{\min\left(N_t,N_r\right)m\sigma_w^2 N_t}\|\mathbf{H}\|_F^2\right)\right]. \tag{3.15}$$

The capacity of a MISO is given in (2.6) [6, 30]. Provided that the system implementation can

Simulation Parameters	Values
BS Type	Macro, Micro, Pico, Femtocell
Power Model Parameters	SOTA 2010
Carrier Frequency	2 GHz
Path Loss Model	3GPP NLOS [31]
Iterations (Number of Channels)	100000
Bandwidth	10 MHz
Operating Temperature	Outdoor:290 K, Indoor: 293.5 K

Table 3.2: *Simulation Parameters.*

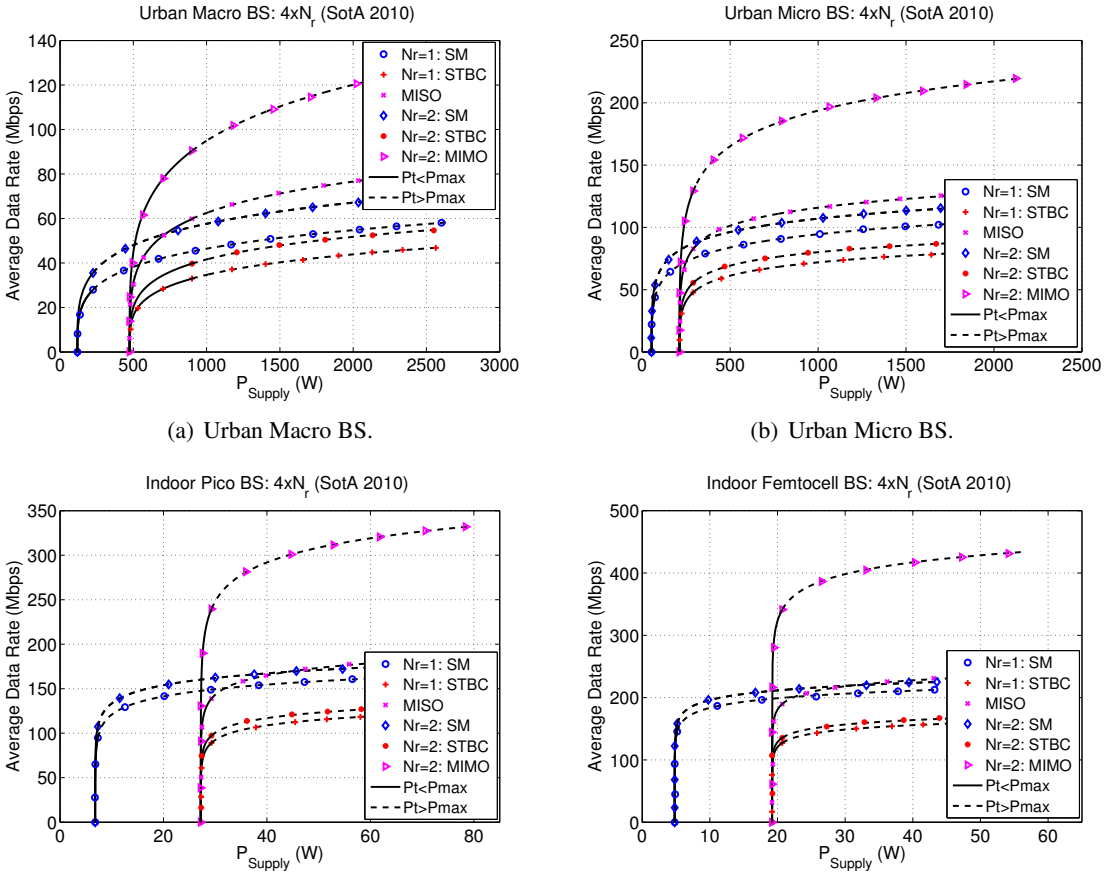
BS type	d_{min} (m)	d_{max} (m)
Macro	10	5000
Micro	10	1000
Pico	10	150
Femtocell	5	50

Table 3.3: *Limits of BS-user distance.*

be done using N_t RF chain, the incorporation of (3.7) in (2.6) gives that:

$$C_{\text{MISO}}\left(P_{\text{supply}}\right) = W \log_2\left(1 + \frac{\left(P_{\text{supply}} - N_t P_0\right) \|\mathbf{h}_{\text{MISO}}\|_2^2}{W m \sigma_w^2}\right). \tag{3.16}$$

In this section, each type of BS is equipped with $N_t = 4$ antennas. Furthermore, it is assumed that the receiver is equipped with one or two antennas. The considered channel model is the 3rd Generation Partnership Project (3GPP) Non-Line of Sight (NLOS) channel model [31]. The 3GPP NLOS channel model generates the wireless channel by taking into account the system setup, propagation environment, and transmission distance. Provided that the system setup and the propagation environment are determined by the studied simulation scenario, the achievable data rates given in (3.10)–(3.16) also depend on the transmission distance between a BS and a user. Therefore, in this thesis, the BS–to–user distance d is selected randomly following a spatially uniform distribution between d_{min} and d_{max}. Also, based on the type of BS, d_{min} and d_{max} are given in Table 3.3. The shadowing standard deviation is 6 dB for the case of urban macro BS and 4 dB for the case of urban micro, indoor pico and femtocell BSs. In addition, the thermal noise power is defined as $N = W\kappa\theta$, where κ is the Boltzmann constant and θ is the operating temperature in K. Based on these assumptions, Monte Carlo simulations are deployed in order to quantify the average data rate and EE of a BS, for the system configuration given in Table 3.2. The calculation of the average data rate of a BS is undertaken following an iterative procedure. During the i-th iteration, the distance d is selected as described before and



(a) Urban Macro BS.

(b) Urban Micro BS.

(c) Indoor Picocell BS.

(d) Indoor Femtocell BS.

Figure 3.4: *Data Rates for different types of BS ($N_t = 4$). The solid lines depict the data rate achieved when the transmission power is restricted by the EARTH power model ($P_t \leq P_{max}$) and the dashed lines depict the data rate achieved when this restriction is not assumed, i.e., $P_t > P_{max}$.*

the data rate $R_i(P_{\mathrm{t}}^{\mathrm{total}})$ is computed. Here, it holds that $P_{\mathrm{t}}^{\mathrm{total}} = N_{RF}P_t$. Finally, the average data rate is written as $\bar{R}(P_{\mathrm{t}}^{\mathrm{total}}) = \mathrm{E}[\mathrm{R_i}(\mathrm{P_t^{total}})]$.

3.4.1 Average Data Rate Results

In Fig. 3.4, the average data rate versus the power supply is depicted, when the BS type is urban macro; urban micro; indoor pico; and indoor femtocell. The illustrated results correspond to two different scenarios of RF power transmission. In the first scenario (solid lines), the RF transmit power is restricted by the EARTH power model ($P_{\mathrm{t}} \leq P_{\mathrm{max}}$), which is a real-world hard constraint. Whereas, in the second scenario (dashed lines), the RF transmit power is increased without any limitation.

Clearly, for $P_{\text{t}} \leq P_{\text{max}}$ (solid lines), in all cases there is a range of rates, where for the same data rate, SM requires significantly less power supply compared to MISO and MIMO (a reduction which can reach up to 76%). However, it can be seen that MISO and MIMO are capable of achieving higher data rates than SM with the significant cost of increasing the already high BS power consumption. As regard to the comparison of SM with STBC, SM achieves higher data rates with significantly less power consumption due to the code dependent fractional rates of STBC.

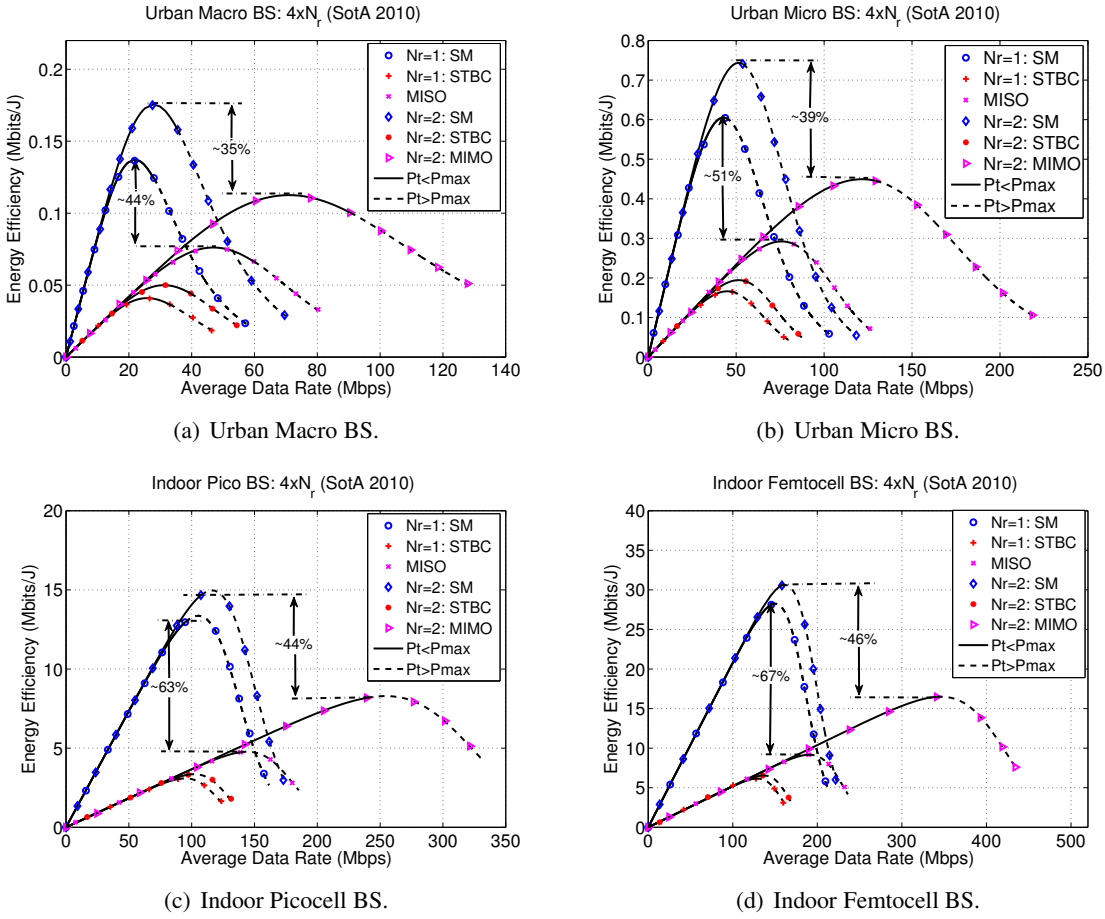
When the RF transmission power is increased beyond the EARTH limit of P_{max} (dashed lines), it can be seen from Fig. 3.4 that there is a crossing point where MISO and MIMO achieve higher data rates with less power consumption. This point happens when the RF transmission power is the dominant consuming factor in the power supply of a BS. For $N_r = 1$, this point occurs at 541, 249, 32, and 23 W for a macro, micro, pico, and femtocell BS, respectively, and for $N_r = 2$, this point occurs at 508, 223, 27, 19 W for a macro, micro, pico, and femtocell BS, respectively. However, these crossing points do not have a practical relevance, since P_{max} is a real-world hard constraint.

3.4.2 Energy Efficiency Results

In this study, the definition of the EE of a BS, in terms of total power consumption, is given as:

$$\text{EE}(P_{\text{t}}^{\text{total}}) = \frac{\bar{R}(P_{\text{t}}^{\text{total}})}{P_{\text{supply}}(P_{\text{t}}^{\text{total}})}. \tag{3.17}$$

The EE performance of different BS types, measured in Mbits/J, under different MIMO transmission schemes is presented in Fig. 3.5. As shown in Fig. 3.5, SM is the most energy efficient transmission scheme for all types of BS, when $P_{\text{t}} \leq P_{\text{max}}$ (solid lines). This finding arises from the fact that SM has a lower constant power consumption due to the single RF chain configuration. In addition, Fig. 3.5 shows that the EE of the studied schemes is maximized for a certain average data rate and power transmission point. Also, in Fig. 3.5, the comparison of the maximum EE of each considered scheme is presented. In particular, in Fig. 3.5, it is shown that, MISO is 67% less energy efficient compared to SM and MIMO is 46% less energy efficient compared to SM. The EE difference between SM and STBC is even higher. Note that the EE difference between SM and the benchmark system is even higher for the rate for



(a) Urban Macro BS.

(b) Urban Micro BS.

(c) Indoor Picocell BS.

(d) Indoor Femtocell BS.

Figure 3.5: *EE Results for different types of BS ($N_t = 4$). The solid lines depict the EE achieved when the transmission power is restricted by the EARTH power model ($P_t \leq P_{max}$) and the dashed lines depict the EE achieved when this restriction is not assumed, i.e., $P_t > P_{max}$.*

which the EE of SM is maximized. This form of comparison is useful when the objective is not the maximization of the spectral efficiency, but the maximization of the EE in term of Mbits/J. Such an example is a scenario where the required data rate is below or equal to the rate for which the EE of SM is maximized and the latency requirement are satisfied both from SM and the benchmark systems. In this case, the main objective is the maximization of the EE. Finally, when the design criterion is the maximization of the EE, the previous EE differences and the rates for which the EE is maximized can be exploited in an adaptive way in order to offer EE in a cellular network without affecting the QoS of the users. Such an example is the scheme presented in [168].

Finally, Fig. 3.5 shows that when the RF transmission power is increased above $P_{\mathrm{t}} > P_{\mathrm{max}}$

(dashed lines), there is a crossing point where MISO and MIMO become more energy efficient than SM. When the mobile terminal is equipped with one receive antenna, this crossing point occurs at 39.18, 72.5, 150.5, and 201.1 Mbps for a macro, micro, pico, and femtocell BS, respectively. And when the mobile terminal is equipped with two receive antennas, this crossing point occurs at 48.25, 82.57, 160.3, and 210.9 Mbps for a macro, micro, pico, and femtocell BS, respectively. As it is stated before, these crossing points do not have a practical value, since they violate the hard constraint of the maximum transmission power ($P_{\text{t}} \leq P_{\text{max}}$).

As regard to the comparison between SM and STBC, it can be seen that SM is always more energy efficient. Hence, taking into account that the current and future predicted hardware technology are restricted by the hard constraint of P_{max}, SM is the most energy efficient scheme among the benchmark systems.

3.5 Summary

In this chapter, using the limits of channel capacity and Monte Carlo simulations, it was demonstrated that the single RF chain configuration of SM enables significant EE gains compared with MISO, MIMO and STBC, under all types of BSs. In more detail, it was shown that for all evaluated multi-antenna schemes and types of BSs there is a transmission point where the EE is maximized. In addition, it was concluded that for a range of data rates, SM offers the same data rate as the benchmark systems with significantly less power consumption, resulting into a reduction which can be as low as 76% for four transmit antennas.

Chapter 4
Point-to-point Receive-Spatial Modulation: Modulating Information at the Receive Antennas

4.1 Introduction

The main objective of this chapter is to evaluate the performance of R-SM and MSR-SM. Initially, the theoretical ABEP of both schemes is derived using the union bound technique when: i) the wireless channel follows the Rayleigh distribution; ii) the precoding method is ZF with P-CSIT. In more detail, an accurate statistical framework which characterizes the received signal is proposed. Using this framework, analytical upper bounds which describe the ABEP performance of R-SM and MSR-SM are developed. In addition, the diversity order and coding gain of R-SM and MSR-SM are derived based on the new bounds. Also, using Monte Carlo simulation results, in Chapter 4, the BER performance of R-SM and MSR-SM is compared against the performance of the corresponding conventional MIMO scheme. In addition, assuming I-CSIT, two regularized linear precoding methods are proposed. The design of the first precoder is based on I-CSIT in a statistical form, whereas the second precoder uses I-CSIT in a worst-case form. Finally, in Chapter 4, a theoretical framework which characterizes the instantaneous and average power of all precoders used in this thesis is proposed.

4.2 Point-to-Point Receive-Spatial Modulation using ZF Precoding and Perfect Channel State Information at the Transmitter

4.2.1 System Model

An uncoded point-to-point MIMO system with N_t transmit antennas and N_r receive antennas is considered. In Chapter 4, the focus is on the downlink where the assumption that the transmitter (base station) is equipped with more antennas than the receiver (mobile terminal) is valid ($N_t \geq$

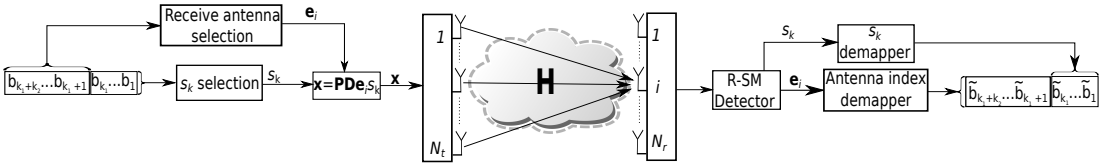


Figure 4.1: *The system model of Receive-Spatial Modulation.*

N_r). Under the assumption that the transmitter employs a linear precoder and the wireless channel is flat quasi-static, the system equation is expressed as in (2.14) as:

$$\mathbf{y} = \mathbf{HPDx} + \mathbf{w}. \tag{4.1}$$

It is clear that the system equation of (4.1) forms a closed-loop SMX scheme which conveys N_r constellation points, if the elements of $\mathbf{x}$ are drawn from a signal constellation $\mathcal{M}$.

In this section, under the assumption of P-CSIT, the focus is on ZF linear precoding. In conventional MIMO schemes, ZF (Bezout) precoding is a suboptimal method that offers a good trade-off between complexity and performance [74]. The main characteristic of ZF is the total elimination of the ICI at the receiver ($\mathbf{HP}_{\text{ZF}} = \mathbf{I}_{N_r,N_r}$). The ZF precoding matrix is just the Moore-Penrose pseudo-inverse of the channel matrix $\mathbf{H}$ and is given in (2.18) when the Singular Value Decomposition (SVD) is used. An alternative form of the ZF precoder is attained in (2.19) when the QR decomposition is used.

Under the assumption of ZF precoding, if (2.19) is plugged in (4.1), the received signal is given as:

$$\mathbf{y} = \mathbf{Dx} + \mathbf{w}. \tag{4.2}$$

Note that when the precoding method is ZF, the elements of the main diagonal of $\mathbf{D}$ take the following form:

$$d_i = \sqrt{\frac{1}{\left[(\mathbf{HH}^H)^{-1}\right]_{i,i}}},\ i = 1, \ldots, N_r, \tag{4.3}$$

where, $[\mathbf{A}]_{i,i}$ is the (i,i) element of matrix $\mathbf{A}$.

In a R-SM communication system, such as the example given in Fig. 4.1, the bitstream to be transmitted during a symbol period is encoded: i) on the index of a single receiving antenna (out of N_r) and ii) in the transmission of a constellation point drawn from a conventional diagram like M-ary QAM or Binary Phase Shift Keying (BPSK). For this reason, the R-SM

transmission mechanism dictates that the transmitted vector $\mathbf{x}$ is written as:

$$\mathbf{x} = \mathbf{e}_i s_k. \tag{4.4}$$

Here, $\mathbf{e}_i$, $i = 1, \ldots, N_r$, is a vector which is equal to the i-th column of $\mathbf{I}_{N_r,N_r}$. Hence, all the elements of $\mathbf{e}_i$ are zero except the one in the i-th position which has the value of one. Furthermore, $s_k \in \mathcal{M} = \{s_1, \ldots, M\}$ is a constellation point which belongs to a M-ary diagram, such as QAM or BPSK.

In this way, if the number of receive antennas is set to be a power of two, the spectral efficiency of R-SM is equal to $k_{\text{R-SM}} = \log_2(M) + \log_2(N_r)$ bpcu. Here, $k_1^{\text{R-SM}} = \log_2(M)$ bits are encoded via the transmission of the constellation point s_k and $k_2^{\text{R-SM}} = \log_2(N_r)$ bits are conveyed via the selection of one (out of N_r) receiving antenna.

Given that the employed precoding method is ZF, at the receiver side the transmitted bitstream can be recovered via the detection of the transmitted constellation point s_k and the index of the receiving antenna. The optimum ML detector is expressed as:

$$(\tilde{\mathbf{x}}) = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2^2. \tag{4.5}$$

As shown in (4.5), the index of the receiving antenna i and the transmitted constellation point s_k are jointly detected via the detection of $\mathbf{x} = \mathbf{e}_i s_k$.

4.2.2 Extension to Multiple Streams

In conventional SM, the main rationale behind the formation of a single stream from the transmitter to the receiver is the deployment of a single RF chain. However, this is not the case for R-SM. Due to the linear precoding at the transmitter, transmission requires multiple RF chains. Thus, the transmission of multiple symbol streams does not impose a significant complexity increase at the transmitter. For this reason, the authors of [21] extend the concept of R-SM to a scheme which establishes multiple symbol streams between the communicating ends. In fact, [21] applies the concept of MS-SM of [19] at the receiver side. In this thesis, the extension of the concept of MS-SM at the receiver side is termed as MSR-SM.

In a wireless system which uses MSR-SM for communication, such as the one in Fig. 4.2, similar to R-SM, the transmitter spatially modulates not one but $N_s \leq N_r$ symbol streams to

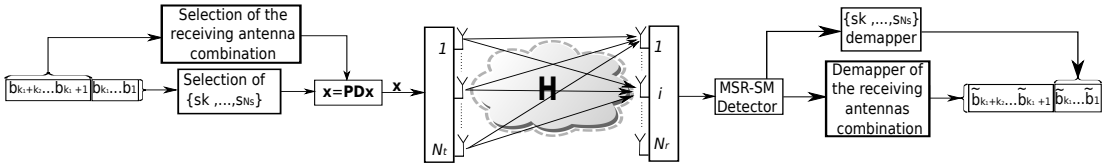


Figure 4.2: *The system model of Multi-Stream Receive-Spatial Modulation.*

the indices of the multiple receiving antennas. The system equation of MSR-SM is given from (4.1).

The objective of MSR-SM is twofold: i) the establishment of $N_s \leq N_r$ parallel symbol streams from the transmitter to the receiver; and ii) additional information conveyance via the indices of the N_s (out of N_r) receiving antennas. Provided that the deployed precoding method is ZF, the received signal is given in (4.2). Hence, by selecting properly the structure of the transmit signal vector $\mathbf{x}$, it possible to enforce that the noise free received signal $\mathbf{Dx}$ has exactly N_s non-zero elements and $N_r - N_s$ zero elements. Given that $\mathbf{D}$ is the diagonal matrix described in the previous section, the non-zero elements of $\mathbf{Dx}$ is a scaled version of the corresponding non-zero elements of $\mathbf{x}$. Similarly, the positions of the zero elements of $\mathbf{Dx}$ are the same as the positions of the zero elements of $\mathbf{x}$. For this reason, a portion of binary information can be encoded on the position of the non-zero elements of $\mathbf{x}$ and consequently on the position of the non-zero elements of the noise free received signal $\mathbf{Dx}$.

In MSR-SM, the structure of $\mathbf{x}$ is given as:

$$\mathbf{x} = \Big[0,\ldots,0,\underbrace{s_1}_{i_1\text{-th position}},0,\ldots,0,\underbrace{s_i}_{i_k\text{-th position}},0,\ldots,0,\underbrace{s_{N_s}}_{i_{N_s}\text{-th position}}0\ldots,0\Big]^T \tag{4.6}$$

where, $\{s_1,\ldots,s_{N_s}\} \in \mathcal{M}$. Here, $\mathcal{M}$ is the deployed constellation. The positions of the non-zero elements correspond to the indices of the receiving antennas, while the positions of the zero elements correspond to the antennas which do not receive signal. Given that the length of $\mathbf{x}$ is N_r and there are N_s non-zero elements, the number of total combinations of N_s non-zero elements (receiving antennas) out of N_r is $\binom{N_r}{N_s}$. However, only $k_c = 2^{\lfloor \log_2\left(\binom{N_r}{N_s}\right)\rfloor}$ combinations are used in order to encode $k_2^{\text{MSR-SM}} = \lfloor \log_2\left(\binom{N_r}{N_s}\right)\rfloor$ bits. This is done by assigning a unique binary index of length of $k_2^{\text{MSR-SM}}$ bits to each one of the k_c selected combinations. The selection of these combinations of receiving antennas can be done intelligently and in adaptive way, based on the values of $d_i, i = 1,\ldots,N_t$, in order to increase the detection (receive) SNR

and minimize the instantaneous BER. Alternatively, it can be done randomly. In the first case, the system complexity is increased and the selection has to be done adaptively. Given that MSR-SM is a closed loop scheme, the latter case is expected to offer a good performance with no further complexity overhead. In this thesis, the focus is on the latter case.

In this way, every symbol period, the bit-stream to be transmitted is divided in two portions. The first portion, which is of length of $k_1^{\text{MSR-SM}} = N_s \log_2(M)$ bits, is encoded and transmitted using N_s symbols drawn from the M-ary constellation $\mathcal{M}$. Provided that each legal combination of N_s receiving antenna encodes $k_2^{\text{MSR-SM}} = \lfloor \log_2 \left(\binom{N_r}{N_s} \right) \rfloor$ bits, the spectral efficiency of MSR-SM is:

$$k_{\text{MSR-SM}} = N_s \log_2(M) + \lfloor \log_2 \left(\binom{N_r}{N_s} \right) \rfloor, \tag{4.7}$$

bpcu.

At the receiver side, the transmitted bit-stream is recovered via the detection of the conventional constellation points $\{s_1, \ldots, s_{N_s}\}$ and the combination of the receiving antennas $\{i_1, \ldots, i_{N_s}\}$. Mathematically, the ML detector of MSR-SM is given from (4.5). However, the alphabet of all possible transmitted symbol vectors $\mathbf{x}$ of MSR-SM is given from (4.6).

4.3 Theoretical Evaluation of the Average Bit Error Probability

In this section, using the union bound technique [169], an upper bound of the ABEP of R-SM and MSR-SM is provided. Based on the union bound technique, the ABEP is upper bounded as:

$$P_{\text{bit}}(\gamma) \leq \frac{1}{|\mathcal{B}| k_t} \sum_{\mathbf{x}} \sum_{\substack{\hat{\mathbf{x}} \\ \hat{\mathbf{x}} \neq \mathbf{x}}} d(\mathbf{x} \rightarrow \hat{\mathbf{x}}) P_{\text{e}}(\mathbf{x} \rightarrow \hat{\mathbf{x}}, \gamma), \tag{4.8}$$

where, $P_{\text{bit}}(\gamma)$ is the ABEP for a given transmit SNR $\gamma = \frac{1}{\sigma_{\mathbf{w}}^2}$ (it is reminded that $\mathrm{E}_{\mathbf{x}}[\mathbf{x}] = 1$). Furthermore, the set of all possible transmitted symbol vectors $\mathbf{x}$ is denoted as $\mathcal{B}$. The size of $\mathcal{B}$ (number of all possible transmitted symbol vectors) is written as $|\mathcal{B}|$. The number of bits transmitted per symbol period is represented by k_t. Also, $P_{\text{e}}(\mathbf{x} \rightarrow \hat{\mathbf{x}}, \gamma)$ denotes the Pairwise Error Probability (PEP) of transmitting the symbol vector $\mathbf{x}$ and detecting the symbol vector $\hat{\mathbf{x}}$. Finally, $d(\mathbf{x} \rightarrow \hat{\mathbf{x}})$ is the Hamming distance which is defined as the number of different bits between the bit-word represented by $\mathbf{x}$ and the bit-word represented by $\hat{\mathbf{x}}$.

The computation of (4.8) requires the evaluation of the probability $P_{\text{e}}(\mathbf{x} \rightarrow \hat{\mathbf{x}}, \gamma)$. In order to

derive the previous PEP, initially, the instantaneous PEP (which is conditioned on the instantaneous channel) has to be obtained. In the next step, the average of the instantaneous PEP over all channel realizations needs to be computed.

Provided that the detection process is performed using (4.5), a symbol error occurs when:

$$\mathcal{E}\left(\mathbf{x},\hat{\mathbf{x}}\right)=\left\{\|\mathbf{y}-\mathbf{D}\mathbf{x}\|_2^2>\|\mathbf{y}-\mathbf{D}\hat{\mathbf{x}}\|_2^2\right\}. \tag{4.9}$$

Using a straightforward elaboration on $\mathcal{E}\left(\mathbf{x},\hat{\mathbf{x}}\right)$, it can be show that:

$$\mathcal{E}\left(\mathbf{x},\hat{\mathbf{x}}\right)=\left\{-\text{Re}\{\mathbf{c}^H\mathbf{D}\mathbf{w}\}>\frac{\mathbf{c}^H\mathbf{D}^2\mathbf{c}}{2}\right\}, \tag{4.10}$$

or

$$\mathcal{E}\left(\mathbf{x},\hat{\mathbf{x}}\right)=\left\{-\sum_{i=1}^{N_r}d_i\text{Re}\{\mathbf{c}_i^*\mathbf{w}_i\}\}>\frac{\mathbf{c}^H\mathbf{D}^2\mathbf{c}}{2}\right\}, \tag{4.11}$$

where, $\mathbf{c}=\mathbf{x}-\hat{\mathbf{x}}$. Also, $\mathbf{c}_i$ and $\mathbf{w}_i$, $i=1,\ldots,N_r$, is the i-th elements of $\mathbf{c}$ and $\mathbf{w}$ respectively. Given that:

$$\text{Re}\{\mathbf{c}_i^*\mathbf{w}_i\}=\text{Re}\{\mathbf{c}_i\}\text{Re}\{\mathbf{w}_i\}+\text{Im}\{\mathbf{c}_i\}\text{Im}\{\mathbf{w}_i\} \tag{4.12}$$

and

$$\text{Re}\{\mathbf{w}_i\}\text{Im}\{\mathbf{w}_i\}\sim\mathcal{N}\left(0,\frac{\sigma_{\mathbf{w}}^2}{2}\right). \tag{4.13}$$

Therefore, by combining (4.10)–(4.13), it can be shown that:

$$-\text{Re}\{\mathbf{c}^H\mathbf{D}\mathbf{w}\}\sim\mathcal{N}\left(0,\sigma_{\mathbf{w}}^2\frac{\mathbf{c}^H\mathbf{D}^2\mathbf{c}}{2}\right). \tag{4.14}$$

Hence, by taking into account the statistical distribution of (4.14) and using the fact that $\gamma=\frac{1}{\sigma_{\mathbf{w}}^2}$, the instantaneous PEP is expressed as:

$$P_{\text{e}}(\mathbf{x}\rightarrow\hat{\mathbf{x}},\gamma|\mathbf{D}^2)=Q\left(\sqrt{\frac{\mathbf{c}^H\mathbf{D}^2\mathbf{c}}{2}\gamma}\right)=Q\left(\sqrt{\frac{z}{2}\gamma}\right), \tag{4.15}$$

where, $z=\mathbf{c}^H\mathbf{D}^2\mathbf{c}$ and $Q\left(\cdot\right)$ is the Q-function.

4.3.1 Receive-Spatial Modulation: Single Stream

In this subsection, the PEP of the single stream R-SM is derived. In this way, using this result, the bound of the ABEP of R-SM is obtained via (4.8). In order to simplify (4.15), the RV z has to be characterized. Thus, the different types of symbol errors in R-SM have to be considered. Similar to SM, R-SM faces three type of symbol errors: i) signal; ii) space; and iii) joint errors. A signal error $\mathcal{E}_1 = \{\mathbf{x} \to \hat{\mathbf{x}} | \{s_k \to \hat{s}_k, \mathbf{e}_i \to \mathbf{e}_i\}\}$ occurs when the index of the receiving antenna is correctly detected and the conventional constellation point s_k is incorrectly detected. In contrast, the incorrect detection of the index of the receiving antenna with the concurrent correct detection of the conventional constellation point s_k results in a space error $\mathcal{E}_2 = \{\mathbf{x} \to \hat{\mathbf{x}} | \{s_k \to s_k, \mathbf{e}_i \to \mathbf{e}_{\hat{i}}\}\}$. Finally, a joint error takes place when both the index of the receiving antenna and the constellation point are incorrectly detected $\mathcal{E}_3 = \{\mathbf{x} \to \hat{\mathbf{x}} | \{s_k \to \hat{s}_k, \mathbf{e}_i \to \mathbf{e}_{\hat{i}}\}\}$.

Based on the different types of symbol errors and using the structure of $\mathbf{x}$ and $\hat{\mathbf{x}}$, the RV z is written as:

$$z = \begin{cases} |s_k - \hat{s}_k|^2 d_i^2, & \text{for } \mathcal{E}_1, \quad (4.16a) \\ |s_k|^2 (d_i^2 + d_{\hat{i}}^2), & \text{for } \mathcal{E}_2, \quad (4.16b) \\ |s_k|^2 d_i^2 + |\hat{s}_k|^2 d_{\hat{i}}^2, & \text{for } \mathcal{E}_3. \quad (4.16c) \end{cases}$$

Considering (4.15) and (4.16a)–(4.16c), it can be seen that the derivation of the average PEP requires the marginal Probability Density Function (PDF) of each d_i^2, $i = 1, \ldots, N_r$, for $\mathcal{E}_1$. In addition, the same equations reveal that derivation of the PEP of $\mathcal{E}_2$ and $\mathcal{E}_3$ can be done using the joint PDF of each pair between d_i^2 and $d_{\hat{i}}^2$.

The PDF of d_i^2 is explicitly derived in [42] as a gamma distribution with $d_i^2 \sim \text{Gamma}(L, 1)$ and $L = N_t - N_r + 1$. Here, Gamma (k, θ) stands for the gamma distribution with shape k and scale θ. Thus, the marginal PDF of each d_i^2 is given as:

$$f_{d_i^2}(x) = \frac{1}{\Gamma(L)} x^{L-1} e^{-x} H_0(x), \quad (4.17)$$

where, $H_0(x)$ is the Heaviside step function, for which it holds that $H_0(x) = 0$ for $x < 0$ and $H_0(x) = 1$ for $x \geq 0$. Furthermore, $\Gamma(\cdot)$ denotes the gamma function defined in [170, p. 892].

The observation of the structure of d_i^2, $i = 1, \ldots, N_r$, in (4.3) shows that they are correlated

gamma RV (with marginal PDF given in (4.17)), as they are produced from the same random matrix $\mathbf{H}$ using the same mathematical formula [1]. Thus, as given in [171, p. 337], the joint PDF of d_i^2 and $d_{\hat{i}}^2$ is the Kibble's bivariate gamma distribution:

$$h_{d_i^2,d_{\hat{i}}^2}(x,y) = \frac{(1-\rho_c)^{-L}}{\Gamma(L)} \sum_{k=0}^{+\infty} b_k f_k(x) p_k(y), \tag{4.18}$$

where,

$$b_k = \frac{\rho_c^k}{(1-\rho_c)^{2k}\Gamma(L+k)k!}, \tag{4.19}$$

$$f_k(x) = \left[x^{L+k-1} e^{-\frac{x}{1-\rho_c}}\right] H_0(x), \tag{4.20}$$

and

$$p_k(y) = \left[y^{L+k-1} e^{-\frac{y}{1-\rho_c}}\right] H_0(y). \tag{4.21}$$

Also, in (4.18), ρ_c stands for the Pearson product-moment correlation coefficient with:

$$\rho_c = \frac{\mathrm{E}\left[(d_i^2 - L)(d_{\hat{i}}^2 - L)\right]}{L}. \tag{4.22}$$

Given the statistical characterization of d_i^2 and $d_{\hat{i}}^2$, the estimation of ρ_c can be carried out, for certain values of N_t and N_r, by using multiple samples of the two RVs.

In order to obtain the average PEP of R-SM, when a signal error $\mathcal{E}_1$ takes place, (4.15) can be simplified using (4.16a). In this way, the average of the simplified version of (4.15), over all possible realizations of d_i^2, is expressed as:

$$P_{\mathcal{E}_1\gamma}(\mathbf{x} \to \hat{\mathbf{x}}, \gamma) = \mathrm{E}_{d_i^2}\left[Q\left(\sqrt{\frac{|\delta|^2 d_i^2}{2}\gamma}\right)\right], \tag{4.23}$$

where, $\delta = s_k - \hat{s}_k$. In this chapter, an upper bound for (4.23) is achieved by considering the fact that the $Q(\cdot)$ function is tightly upper-bounded as [172]:

$$Q(x) \leq \frac{1}{6}e^{-2x^2} + \frac{1}{12}e^{-x^2} + \frac{1}{4}e^{\frac{-x^2}{2}}. \tag{4.24}$$

[1] Additional justifications are provided in Chapter 5.4, where a more general scenario is considered.

Hence, (4.23) becomes:

$$P_{\mathcal{E}_1}(\mathbf{x} \to \hat{\mathbf{x}}, \gamma) \leq \mathrm{E}_{d_i^2}\left[\frac{1}{6}e^{|\delta|^2 d_i^2 \gamma}\right] + \mathrm{E}_{d_i^2}\left[\frac{1}{12}e^{\frac{|\delta|^2 d_i^2}{2}\gamma}\right] + \mathrm{E}_{d_i^2}\left[\frac{1}{4}e^{\frac{|\delta|^2 d_i^2}{4}\gamma}\right]. \tag{4.25}$$

The right side of (4.25) shows that it requires the evaluation of three expectations which have the form of:

$$g_1(\alpha_1, \beta_1) = \mathrm{E}_{d_i^2}\left[\alpha_1 e^{\beta_1 d_i^2}\right]. \tag{4.26}$$

Here, α_1 and β_1 are positive real constants which depend on each term of the right side of (4.25). Thus, using the marginal PDF of (4.18), it can be shown that:

$$g_1(\alpha_1, \beta_1) = \frac{\alpha_1}{\Gamma(L)} \int_0^{+\infty} x^{L-1} e^{(\beta_1+1)x} = \alpha_1 \left(\beta_1 + 1\right)^{-L}, \tag{4.27}$$

where, in the last step of (4.27), the integration formula from [170, p. 892, 3.351, 3] is considered. Finally, the incorporation of (4.27) in (4.25) gives that:

$$P_{\mathcal{E}_1}(\mathbf{x} \to \hat{\mathbf{x}}, \gamma) \leq \frac{1}{6}\left(|\delta|^2\gamma + 1\right)^{-L} + \frac{1}{12}\left(\frac{|\delta|^2\gamma}{2} + 1\right)^{-L} + \frac{1}{12}\left(\frac{|\delta|^2\gamma}{2} + 1\right)^{-L}. \tag{4.28}$$

The average PEP of R-SM when a space error $\mathcal{E}_2$ occurs can be bounded by following a similar procedure as before. After the simplification of (4.15) using (4.16b), the previous bound of the Q-function can be applied in order to obtain:

$$\begin{aligned} P_{\mathcal{E}_2}(\mathbf{x} \to \hat{\mathbf{x}}, \gamma) \leq{}& \mathrm{E}_{d_i^2, d_{\hat{i}}^2}\left[\frac{1}{6}e^{-|s_k|^2\left(d_i^2 + d_{\hat{i}}^2\right)\gamma}\right] + \mathrm{E}_{d_i^2, d_{\hat{i}}^2}\left[\frac{1}{12}e^{-\frac{|s_k|^2\left(d_i^2 + d_{\hat{i}}^2\right)\gamma}{2}}\right] \\ &+ \mathrm{E}_{d_i^2, d_{\hat{i}}^2}\left[\frac{1}{4}e^{-\frac{|s_k|^2\left(d_i^2 + d_{\hat{i}}^2\right)\gamma}{4}}\right]. \end{aligned} \tag{4.29}$$

The observation of (4.29) shows that the expectations that need to be evaluated have the form of:

$$g_2(\alpha_2, \beta_2) = \mathrm{E}_{d_i^2, d_{\hat{i}}^2}\left[\alpha_2 e^{-\beta_2\left(d_i^2 + d_{\hat{i}}\right)}\right], \tag{4.30}$$

where again α_2 and β_2 are positive real scalars which are defined from (4.29). Given that the joint PDF of the RVs d_i^2 and $d_{\hat{i}}^2$ is given in (4.18), after some straightforward manipulations, it is shown that:

$$g_2(\alpha_2, \beta_2) = \frac{\alpha_2 \left(1 - \rho_c\right)^L}{\Gamma\left(L\right)} \sum_{k=0}^{+\infty} \left[\int_0^{+\infty} e^{-\beta_2 x} f_k(x) \mathrm{d}x \int_0^{+\infty} e^{-\beta_2 y} p_k(y) \mathrm{d}y\right]. \tag{4.31}$$

If the structure of $f_k(x)$ and $p_k(y)$ is taken into account, both integrations in the infinite summation of (4.31) can be reduced to:

$$\int_0^{+\infty} e^{-\beta_2 x} f_k(x) \mathrm{d}x = (L+k-1)! \left(\beta_2 + \frac{1}{1-\rho_c}\right)^{-L-k} \tag{4.32}$$

$$\int_0^{+\infty} e^{-\beta_2 y} p_k(y) \mathrm{d}y = (L+k-1)! \left(\beta_2 + \frac{1}{1-\rho_c}\right)^{-L-k}. \tag{4.33}$$

In the derivation of (4.32) and (4.33), the formula from [170, p. 892, 3.351, 3] is deployed. Using (4.32) and (4.33) in (4.31), after some rearrangements and simplifications, it can be shown that:

$$g_2(\alpha_2, \beta_2) = \frac{\alpha_2 \left(1-\rho_c\right)^{-L}}{\tau\left(\beta_2\right)^{2L}} \, {}_1F_0\left(L; \, ; \frac{\rho_c \tau \left(\beta_2\right)^{-2}}{\left(1-\rho_c\right)^2}\right), \tag{4.34}$$

where,

$$\tau\left(\beta_2\right) = \beta_2 + \frac{1}{1-\rho_c} \tag{4.35}$$

and

$${}_1F_0\left(L; \, ; y\right) = \sum_{k=0}^{+\infty} \frac{[L]_k}{k!} y^k \tag{4.36}$$

is the generalized hypergeometric function, as defined in [170, p. 1010]. Furthermore,

$$[L]_k = L(L+1)\ldots(L+k-1) \tag{4.37}$$

is the Pochhammer symbol. Finally, the incorporation of (4.34) in (4.29) gives:

$$\begin{aligned} P_{\mathcal{E}_2}(\mathbf{x} \to \hat{\mathbf{x}}, \gamma) \leq & \frac{\left(1-\rho_c\right)^{-L}}{6\tau\left(|s_k|^2\gamma\right)^{2L}} \, {}_1F_0\left(L; \, ; \frac{\rho_c \tau \left(|s_k|^2\gamma\right)^{-2}}{\left(1-\rho_c\right)^2}\right) \\ & + \frac{\left(1-\rho_c\right)^{-L}}{12\tau\left(\frac{|s_k|^2\gamma}{2}\right)^{2L}} \, {}_1F_0\left(L; \, ; \frac{\rho_c \tau \left(\frac{|s_k|^2\gamma}{2}\right)^{-2}}{\left(1-\rho_c\right)^2}\right) \\ & + \frac{\left(1-\rho_c\right)^{-L}}{4\tau\left(\frac{|s_k|^2\gamma}{4}\right)^{2L}} \, {}_1F_0\left(L; \, ; \frac{\rho_c \tau \left(\frac{|s_k|^2\gamma}{4}\right)^{-2}}{\left(1-\rho_c\right)^2}\right). \end{aligned} \tag{4.38}$$

The final step of the provided proof is the derivation of the bound of the average PEP of R-SM, when a joint symbol error $\mathcal{E}_3$ happens. Similar to the other types of symbol errors in R-SM, the

combination (4.15) and (4.16c) results in:

$$P_{\mathcal{E}_3}(\mathbf{x}\to\hat{\mathbf{x}},\gamma) \leq \mathrm{E}_{d_i^2,d_{\hat{i}}^2}\left[\frac{1}{6}e^{-\left(|s_k|^2 d_i^2+|\hat{s}_k|^2 d_{\hat{i}}^2\right)\gamma}\right] + \mathrm{E}_{d_i^2,d_{\hat{i}}^2}\left[\frac{1}{12}e^{-\frac{\left(|s_k|^2 d_i^2+|\hat{s}_k|^2 d_{\hat{i}}^2\right)\gamma}{2}}\right] + \mathrm{E}_{d_i^2,d_{\hat{i}}^2}\left[\frac{1}{4}e^{-\frac{\left(|s_k|^2 d_i^2+|\hat{s}_k|^2 d_{\hat{i}}^2\right)\gamma}{4}}\right]. \tag{4.39}$$

The observation of (4.39) shows that it requires the assessment of three terms of the form of:

$$g_3(\alpha_3,\beta_3) = \mathrm{E}_{d_i^2,d_{\hat{i}}^2}\left[\alpha_3 e^{-\beta_3\left(|s_k|^2 d_i^2+|\hat{s}_k|^2 d_{\hat{i}}^2\right)\gamma}\right]. \tag{4.40}$$

Here, the real and positive values of α_3 and β_3 are defined from the right side of (4.38).

The derivation of $g_3(\alpha_3,\beta_3)$ can be obtained by following similar steps as (4.31)–(4.34):

$$\begin{aligned} g_3(\alpha_3,\beta_3) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \alpha_3 e^{-\beta_3\left(|s_k|^2 d_i^2+|\hat{s}_k|^2 d_{\hat{i}}^2\right)\gamma} h_{d_i^2,d_{\hat{i}}^2}(x,y)\mathrm{d}x\mathrm{d}y \\ &= \frac{\alpha_3\left(1-\rho_c\right)^{-L}}{\left(\tau_1\left(\beta_3\right)\tau_2\left(\beta_3\right)\right)^L}\,{}_1F_0\left(L;\,;\frac{\rho_c}{\left(1-\rho_c\right)^2\tau_1\left(\beta_3\right)\tau_2\left(\beta_3\right)}\right), \end{aligned} \tag{4.41}$$

where,

$$\tau_1\left(\beta_3\right) = \beta_3|s_k|^2\gamma + \frac{1}{1-\rho_c} \tag{4.42}$$

and

$$\tau_2\left(\beta_3\right) = \beta_3|\hat{s}_k|^2\gamma + \frac{1}{1-\rho_c}. \tag{4.43}$$

In this way, using the incorporation of (4.41) in (4.39), it can be shown that the average PEP of a joint error is bounded by

$$\begin{aligned} P_{\mathcal{E}_3}(\mathbf{x}\to\hat{\mathbf{x}},\gamma) \leq\ & \frac{\left(1-\rho_c\right)^{-L}}{6\left(\tau_1\left(1\right)\tau_2\left(1\right)\right)^L}\,{}_1F_0\left(L;\,;\frac{\rho_c}{\left(1-\rho_c\right)^2\tau_1\left(1\right)\tau_2\left(1\right)}\right) \\ &+ \frac{\left(1-\rho_c\right)^{-L}}{12\left(\tau_1\left(\frac{1}{2}\right)\tau_2\left(\frac{1}{2}\right)\right)^L}\,{}_1F_0\left(L;\,;\frac{\rho_c}{\left(1-\rho_c\right)^2\tau_1\left(\frac{1}{2}\right)\tau_2\left(\frac{1}{2}\right)}\right) \\ &+ \frac{\left(1-\rho_c\right)^{-L}}{4\left(\tau_1\left(\frac{1}{4}\right)\tau_2\left(\frac{1}{4}\right)\right)^L}\,{}_1F_0\left(L;\,;\frac{\rho_c}{\left(1-\rho_c\right)^2\tau_1\left(\frac{1}{4}\right)\tau_2\left(\frac{1}{4}\right)}\right). \end{aligned} \tag{4.44}$$

Given that PEPs of R-SM for the different types of symbol error are given from (4.28), (4.38),

and (4.44), the derivation of the upper bound of ABEP of R-SM can be obtained from (4.8) by setting $|\mathcal{B}| = MN_r$ and $k_t = \log_2(M) + \log_2(N_r)$.

4.3.2 Multi-Stream Receive-Spatial Modulation: Multiple Streams

The ABEP performance analysis of MSR-SM is provided (described) in this subsection. The derivation of the ABEP of MSR-SM can be based on (4.8) by setting $|\mathcal{B}| = M^{N_s} 2^{\left\lfloor \log_2 \binom{N_r}{N_s} \right\rfloor}$ and $k_t = N_s \log_2(M) + \lfloor \log_2 \left(\binom{N_r}{N_s} \right) \rfloor$. From (4.8), it can be seen that the assessment of the ABEP of MSR-SM requires the evaluation of its average PEP. This is done by averaging the instantaneous PEP over all possible channel realizations. The derivation of the average PEP is omitted since it can be directly obtained from the analysis presented in Chapter 5.4 by considering only point-to-point communication (single user transmission) and neglecting the large scale channel effect. In particular, the average PEP of the point-to-point MSR-SM is the direct result of (5.44) when a single user communication is assumed and there is no large scale channel effect.

4.4 Diversity Order and Coding Gain of Multi-Stream Receive-Spatial Modulation

In order to gain a better understanding of the performance of MSR-SM, its diversity order and coding gain are obtained. The definition of the diversity order and coding gain is given in [173]. In more detail, the diversity order is the slope of the curve of the Symbol Error Rate (SER) versus SNR, when the SER is expressed in logarithmic scale and the SNR is expressed in dB, and as the SNR approaches infinity ($\gamma \to +\infty$) [173]. Furthermore, the coding gain determines how far to the left to the SER curve is shifted with respect to the benchmark SER which has a coding gain equal to one. However, in most cases and due to its mathematical difficulty, the diversity order cannot be obtained directly from the curve of SER. Therefore, the diversity order and coding gain are obtained from a high SNR approximation of the SER, when $\gamma \to +\infty$ [173]. Mathematically, this is expressed as:

$$\mathrm{SER}^{+\infty} \lessapprox \left(c_{\mathrm{g}} \gamma\right)^{-d_{\mathrm{o}}}, \tag{4.45}$$

where, $\mathrm{SER}^{+\infty}$ is a high SNR approximation of the SER, d_{o} is the diversity order, and c_{g} is the coding gain.

Similar to Section 4.3.2, the diversity order and coding gain of point-to-point MSR-SM are directly obtained from (5.52) and (5.53) in Chapter 5.5, respectively, if single user communication is assumed and the wireless channel experiences only small scale fading. More specifically, a high SNR approximation of the average PEP can be obtained by using the Chernoff bound of the $\mathrm{Q}(\cdot)$ function. In the next step, a high SNR approximation of the average SER is formed by incorporating the previous high SNR approximation of the average PEP into (4.8) and using the same approach as Chapter 5.5. In this way, a high SNR approximation of the average SER is obtained which clearly shows the diversity order coding gain of MSR-SM. The previous proof is omitted since it closely resembles to the proof presented in Chapter 5.5.

Therefore, the diversity order of MSR-SM is:

$$d_{\text{MSR-SM}} = L, \tag{4.46}$$

and its corresponding coding gain is:

$$c_{\text{MSR-SM}} = \frac{d_{\text{min}}^2}{4} \sqrt[L]{\frac{2}{N_s \left(M^2 - M\right)}}. \tag{4.47}$$

In (4.47), d_{min}^2 denotes the minimum distance between every pair of the deployed constellation points. In Chapter 5.5, there is a discussion with respect to the BER performance comparison between MSR-SM and the benchmark system of conventional MIMO (SMX using ZF precoding). In more detail, it is shown that MSR-SM achieves the same diversity order and higher coding gain. The description of this analysis is omitted here since it follows the same arguments as those given in Chapter 5.5

4.5 Imperfect Channel State Information at the Transmitter

In practice, CSI is acquired by the receiver using a pilot sequence. Under the assumption that the channel does not change during one block of symbols, CSIR may be considered accurate. When the channel varies rapidly, even CSIR is subjected to imperfections. At the transmitter side, CSI is acquired either using a low rate feedback from the receiver or using the reciprocity

principle when it is applicable (pure TDD). In the general case, channel imperfections at the transmitter side may occur because of: i) channel estimation errors; ii) quantization errors when a reverse feedback link is employed; and iii) an outdated version of CSIT (time varying channel).

In practical scenarios, the acquisition of perfect (accurate) CSI is either a very expensive or even an unrealistic process. Hence, there is a high need for designs that take into account channel imperfections. A widely used channel model, that captures channel imperfections, expresses the channel matrix as [174]:

$$\mathbf{H} = \bar{\mathbf{H}} + \widetilde{\mathbf{H}}. \tag{4.48}$$

In (4.48), $\bar{\mathbf{H}}$ represents the long term channel evolution which can be accurately acquired by the transmitter. In addition, $\widetilde{\mathbf{H}}$ denotes the channel for which only some kind of statistical or worst-case knowledge is considered possible. A common assumption is to model the long term channel evolution as $\bar{\mathbf{H}} \sim \mathcal{CN}\left(\mathbf{0}, \sigma^2_{\bar{\mathbf{H}}}\mathbf{I}\right)$, while $\widetilde{\mathbf{H}} \sim \mathcal{CN}\left(\mathbf{0}, \sigma^2_{\widetilde{\mathbf{H}}}\mathbf{I}\right)$. Note that each element of $\bar{\mathbf{H}}$ remains constant over a block of symbols. However, the elements of $\widetilde{\mathbf{H}}$ change every symbol period. The instantaneous channel matrix is denoted as $\mathbf{H}$. Finally, in order to avoid any power amplification due to the wireless channel, it is assumed that $\sigma^2_{\bar{\mathbf{H}}} + \sigma^2_{\widetilde{\mathbf{H}}} = 1$.

In the following two subsections, the precoding matrix $\mathbf{P}$ of (4.1) is designed using two methods from optimization theory. These methods are Stochastic Robust Approximation (SRA) and Worst-Case Robust Approximation (WCRA) [175]. The method of SRA follows a statistical approach and solves the resulting minimization problems over their expectation. In contrast, WCRA is a worst-case method and solves the same minimization problems assuming that $\|\widetilde{\mathbf{H}}\|_2 \leq \alpha$. Note that both approaches employ I-CSIT in a statistical or scalar form which does not change rapidly over time.

4.5.1 Objective Function

The objective function that has to be minimized in order to design the precoding matrix $\mathbf{P}$ using I-CSIT is given below. When P-CSIT is available, the ZF precoding matrix of (2.19) is just the pseudo-inverse matrix of $\mathbf{H}$. An alternative way to obtain the ZF precoder is to solve the following N_r minimization problems:

$$\min_{\mathbf{p}_l} \|\mathbf{H}\mathbf{p}_l - \mathbf{e}_l\|_2^2, \ \forall l = 1, ..., N_r, \tag{4.49}$$

where, $\mathbf{e}_l$ is the l-th column of the identity matrix $\mathbf{I}_{N_r,N_r} = [\mathbf{e}_1, ..., \mathbf{e}_{N_r}]$.

4.5.2 ZF-Like Precoding Based on Stochastic Robust Approximation

When SRA is employed, the minimization problems of (4.49) are solved over their expectation [175]:

$$\min_{\mathbf{p}_l} \mathrm{E}\left[\|\mathbf{H}\mathbf{p}_l - \mathbf{e}_l\|_2^2\right], \ \forall l = 1, \ldots, N_r. \tag{4.50}$$

Using the fact that the MIMO channel can be written as in (4.48), where $\mathrm{E}[\widetilde{\mathbf{H}}] = \mathbf{0}$, (4.50) can be reformulated as:

$$\min_{\mathbf{p}_l} \left\{ \|\bar{\mathbf{H}}\mathbf{p}_l - \mathbf{e}_l\|_2^2 + \mathbf{p}_l^H \mathrm{E}\left[\widetilde{\mathbf{H}}^H \widetilde{\mathbf{H}}\right] \mathbf{p}_l \right\}, \tag{4.51}$$

which is obviously a convex optimization problem (sum of quadratic functions). Thus, the solution can be achieved by setting its gradient equal to zero. In this case, the analytical solution is written as:

$$\mathbf{p}_l = \left[\bar{\mathbf{H}}^H \bar{\mathbf{H}} + \mathrm{E}\left[\widetilde{\mathbf{H}}^H \widetilde{\mathbf{H}}\right]\right]^{-1} \bar{\mathbf{H}}^H \mathbf{e}_l, \tag{4.52}$$

or

$$\mathbf{P}_{\text{SRA}} = \left[\bar{\mathbf{H}}^H \bar{\mathbf{H}} + \mathrm{E}\left[\widetilde{\mathbf{H}}^H \widetilde{\mathbf{H}}\right]\right]^{-1} \bar{\mathbf{H}}^H, \tag{4.53}$$

in a collective matrix form. A careful look in $\mathrm{E}[\widetilde{\mathbf{H}}^H \widetilde{\mathbf{H}}]$ reveals its structured form. For example, when $\widetilde{\mathbf{H}}$ is uncorrelated, it holds that:

$$\mathrm{E}[\widetilde{\mathbf{H}}^H \widetilde{\mathbf{H}}] = N_r \sigma_{\widetilde{\mathbf{H}}}^2 \mathbf{I}_{N_t,N_t}. \tag{4.54}$$

Thus, the computation of the previous quantity requires only the knowledge of $\sigma_{\widetilde{\mathbf{H}}}^2$, which is scalar.

4.5.3 ZF-Like Precoding Based on Worst-Case Robust Approximation

When WCRA is utilized, the MIMO wireless channel is expressed as the non-empty and bounded set $\Phi \subseteq \mathcal{C}^{N_r,N_t}$. Here, Φ represents all the possible values of the channel matrix $\mathbf{H}$. Given a feasible precoding vector $\mathbf{p}_l$, the worst case error can be formulated as:

$$e_{\text{wc}}(\mathbf{p}_l) = \sup\left[\|\mathbf{H}\mathbf{p}_l - \mathbf{e}_l\|_2 | \mathbf{H} \in \Phi\right]. \tag{4.55}$$

The aim of this subsection is to design a precoder that minimizes the worst case error $e_{\mathrm{wc}}(\mathbf{p}_l)$. Under this objective, the minimization problem can be formulated as:

$$\min_{\mathbf{p}_l} \sup \left[\|\mathbf{H}\mathbf{p}_l - \mathbf{e}_l\|_2 | \mathbf{H} \in \Phi \right] \ \forall l = 1, \ldots, N_r. \tag{4.56}$$

In this thesis, the Norm Bound Error (NBE) method from [175] is used. In NBE, the uncertainty of $\widetilde{\mathbf{H}}$ is considered within a norm ball of radius α. Thus, the set Φ is written as:

$$\Phi = \left\{ \mathbf{H} = \bar{\mathbf{H}} + \widetilde{\mathbf{H}} | \|\widetilde{\mathbf{H}}\|_2 \leq \alpha \right\}, \tag{4.57}$$

where, $\alpha > 0$. Let,

$$e_{\mathrm{wc}}^{\mathrm{NBE}}(\mathbf{p}_l) = \sup\{\|\bar{\mathbf{H}}\mathbf{p}_l - \mathbf{e}_l + \widetilde{\mathbf{H}}\mathbf{p}_l\|_2| \ \ \|\widetilde{\mathbf{H}}\|_2 \leq \alpha\}, \tag{4.58}$$

be the worst-case error given the precoding vector $\mathbf{p}_l$. After some arithmetic manipulations, it is shown that $e_{\mathrm{wc}}^{\mathrm{NBE}}(\mathbf{p}_l)$ is equal to:

$$e_{\mathrm{wc}}^{\mathrm{NBE}}(\mathbf{p}_l) = \|\bar{\mathbf{H}}\mathbf{p}_l - \mathbf{e}_l\|_2 + \alpha\|\mathbf{p}_l\|_2, \tag{4.59}$$

and it is attained for $\widetilde{\mathbf{H}} = \alpha\mathbf{u}\mathbf{v}^H$, where,

$$\mathbf{u} = \frac{\bar{\mathbf{H}}\mathbf{p}_l - \mathbf{e}_l}{\|\bar{\mathbf{H}}\mathbf{p}_l - \mathbf{e}_l\|_2} \text{ and } \mathbf{v} = \frac{\mathbf{p}_l}{\|\mathbf{p}_l\|_2}, \tag{4.60}$$

given that $\bar{\mathbf{H}}\mathbf{p}_l - \mathbf{e}_l \neq \mathbf{0}$ and $\mathbf{p}_l \neq \mathbf{0}$. Thus, the minimization problem of (4.56) can be reformulated as:

$$\min_{\mathbf{p}_l} \|\bar{\mathbf{H}}\mathbf{p}_l - \mathbf{e}_l\|_2 + \alpha\|\mathbf{p}_l\|_2. \tag{4.61}$$

The minimization problems of (4.61) can be transformed to:

$$\begin{aligned} &\min_{\{t_1,t_2\}} \quad t_1 + at_2 \\ &\text{subject to } \ \|\bar{\mathbf{H}}\mathbf{p}_l - \mathbf{e}_l\|_2 \leq t_1, \ \|\mathbf{p}_l\|_2 \leq t_2, \end{aligned} \tag{4.62}$$

which is solved as a Second Order Cone Programming (SOCP) problem using the interior point method [175]. Alternatively, (4.61) can be transformed in a Tikhonov Regularization (TR)

form:

$$\min_{\mathbf{p}_l} \|\bar{\mathbf{H}}\mathbf{p}_l - \mathbf{e}_l\|_2^2 + \beta\|\mathbf{p}_l\|_2^2, \tag{4.63}$$

for some value of β [175].

Again, (4.61) is convex because $e_{\text{wc}}^{\text{NBE}}(\mathbf{p}_l)$ is the sum of quadratic functions. Thus, the solution can be reached using the gradient condition and is written as:

$$\mathbf{p}_l = \left[\bar{\mathbf{H}}^H\bar{\mathbf{H}} + \beta\mathbf{I}\right]^{-1}\bar{\mathbf{H}}^H\mathbf{e}_l \;\forall l = 1, \ldots, N_r. \tag{4.64}$$

The collective matrix form of the WCRA precoder is expressed as:

$$\mathbf{P}_{\text{WCRA}} = \left[\bar{\mathbf{H}}^H\bar{\mathbf{H}} + \beta\mathbf{I}\right]^{-1}\bar{\mathbf{H}}^H. \tag{4.65}$$

4.5.4 Precoding in the Presence of Transmit and Receive Space Correlation

The problems of (4.50) and (4.63) are different forms of TR [175]. A valuable property of TR theory is that it does not pose any rank restriction on the involved matrices $\bar{\mathbf{H}}$ and $\widetilde{\mathbf{H}}$, as long as the matrices $\bar{\mathbf{H}}^H\bar{\mathbf{H}} + \mathrm{E}\left[\widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}}\right]$ and $\bar{\mathbf{H}}^H\bar{\mathbf{H}} + \beta\mathbf{I}$ are positive definite [175]. Thus, the analytical solutions of (4.53) and (4.65) may enjoy the additional merit of being applicable to spatially correlated channels.

In this thesis, the Kronecker correlation model is employed [176]. According to this correlation model, the MIMO channel can be rewritten as:

$$\mathbf{H} = \mathbf{R}_R^{1/2}\mathbf{H}_w(\mathbf{R}_T^{1/2})^T, \tag{4.66}$$

where, $\mathbf{H}_w \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$. Here, $\mathbf{R}_T$ and $\mathbf{R}_R$ represent the transmit and receive spatial correlation matrices, respectively. In this case, the wireless channel is distributed as $\mathbf{H} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_T \otimes \mathbf{R}_R)$. Usually, the entries of the spatial correlation matrices $\mathbf{R}_R$ and $\mathbf{R}_T$ are generated using an exponential model with $R_T(i,j) = \rho_t^{|i-j|}$ and $R_R(i,j) = \rho_r^{|i-j|}$, where $0 \leq \rho_t, \rho_r \leq 1$. Values of ρ_t and ρ_r close to 0 mean low correlation, whereas values close to 1 mean high correlations.

If the Kronecker correlation model of (4.66) is combined with the model of I-CSIT described

in Section 4.5, the MIMO channel is expressed as:

$$\mathbf{H} = \mathbf{R}_R^{1/2} \left(\bar{\mathbf{H}} + \widetilde{\mathbf{H}} \right) (\mathbf{R}_T^{1/2})^T. \tag{4.67}$$

Here, the matrix $\mathbf{R}_R^{1/2}\bar{\mathbf{H}}(\mathbf{R}_T^{1/2})^T$ represents the fully known part of the channel and the matrix $\mathbf{R}_R^{1/2}\widetilde{\mathbf{H}}(\mathbf{R}_T^{1/2})^T$ represents the channel uncertainty.

Looking at (4.53) and (4.65), it is revealed that only the analytical form of the design based on SRA is affected by the correlated channel. This is because SRA requires the computation of $\mathrm{E}[\widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}}]$. After some arithmetic manipulations this quantity is expressed as:

$$\mathrm{E}[\widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}}] = \sigma^2_{\widetilde{\mathbf{H}}}\mathrm{tr}(\mathbf{R}_R)\mathbf{R}_T^{1/2}(\mathbf{R}_T^{1/2})^T. \tag{4.68}$$

4.6 Transmit Power Analysis

The focus in this subsection is on the transmitter and on its transmission power. Provided that $\mathrm{E}_{\mathbf{x}}\left[\|\mathbf{x}\|_2^2\right] = 1$, the instantaneous transmission power $\|\mathbf{PDx}\|_2^2$ is fully characterized by the quantity $P_{\mathrm{s}} = \|\mathbf{PD}\|_2^2$.

In (4.1), the normalization matrix $\mathbf{D}$ is employed in order to make sure that the transmission power is not amplified by the precoder $\mathbf{P}$ in adverse channel conditions. As shown below, the proposed precoders have a structure that imposes an indirect transmit power constraint. The practical value of this constraint is that it allows to set $\mathbf{D} = \mathbf{I}_{N_r,N_r}$, without amplifying the power transmission to an unacceptable level. Additionally, it can be assumed that $\mathbf{HP} \approx \mathbf{I}_{N_r,N_r}$ (when $\mathbf{P}$ is the MMSE, or SRA, or WCRA precoder) in order to keep the complexity of the receiver low. In this way, there is no need for d_i to be known at the receiver. Hence, the detection algorithm of (4.5) is further simplified by setting $\mathbf{D} = \mathbf{I}_{N_r,N_r}$. *However, note that in the BER simulation results of Section 4.7, the precoding designs of SRA and WCRA are normalized as described in Section 4.2.1. This is done for the sake of fair comparison with the other precoding methods (same transmission power).*

4.6.1 Instantaneous Power Transmission

The precoding methods of SRA and WCRA are forms of TR when I-CSIT is available. As a consequence, it is expected from both designs to offer a reduced power transmission [175]. The

proof begins by taking the SVD of $\bar{\mathbf{H}}$:

$$\bar{\mathbf{H}} = \bar{\mathbf{U}}\bar{\mathbf{\Sigma}}\bar{\mathbf{V}}^H. \tag{4.69}$$

Here, $\bar{\mathbf{U}}$ and $\bar{\mathbf{V}}$ are unitary matrices and $\bar{\mathbf{\Sigma}}$ is a diagonal matrix containing at its main diagonal the N_r singular values $\bar{\sigma}_i$ of $\bar{\mathbf{H}}$ (without loss of generality it is assumed that $\bar{\mathbf{H}}$ has a rank of N_r). If (4.69) is plugged into (4.53), after some arithmetic manipulations the SRA precoder is written as:

$$\mathbf{P}_{\text{SRA}} = \bar{\mathbf{V}}\left[\bar{\mathbf{D}} + N_r\sigma^2_{\widetilde{\mathbf{H}}}\mathbf{I}_{N_t,N_t}\right]^{-1}\bar{\mathbf{\Sigma}}^T\bar{\mathbf{U}}^H, \tag{4.70}$$

where, $\bar{\mathbf{D}}$ is a $N_t \times N_t$ diagonal matrix equal to $\bar{\mathbf{D}} = \mathbf{diag}\left(\bar{\sigma}^2_1, \ldots, \bar{\sigma}^2_{N_r}\right)$. Here, $\bar{\sigma}^2_i$, $i = 1, \ldots, N_r$ are the eigenvalues of $\bar{\mathbf{H}}^H\bar{\mathbf{H}}$.

Taking into account that both $\bar{\mathbf{D}} + N_r\sigma^2_{\widetilde{\mathbf{H}}}\mathbf{I}_{N_t,N_t}$ and $\bar{\mathbf{\Sigma}}$ are diagonal matrices, (4.70) can be further simplified and written as:

$$\mathbf{P}_{\text{SRA}} = \bar{\mathbf{V}}\bar{\mathbf{\Sigma}}_{\text{SRA}}\bar{\mathbf{U}}^H, \tag{4.71}$$

where,

$$\mathbf{\Sigma}_{\text{SRA}} = \mathbf{diag}\left(\frac{\bar{\sigma}_1}{\bar{\sigma}^2_1 + N_r\sigma^2_{\widetilde{\mathbf{H}}}}, \ldots, \frac{\bar{\sigma}_{N_r}}{\bar{\sigma}^2_{N_r} + N_r\sigma^2_{\widetilde{\mathbf{H}}}}\right). \tag{4.72}$$

In order to evaluate the transmit power of the SRA precoder, the Frobenius norm of the precoding matrix of (4.71) has to be computed as:

$$\|\mathbf{P}_{\text{SRA}}\|^2_F = \text{tr}\left(\mathbf{P}^H_{\text{SRA}}\mathbf{P}_{\text{SRA}}\right) = \text{tr}\left(\bar{\mathbf{U}}\mathbf{\Sigma}^T_{\text{SRA}}\mathbf{\Sigma}_{\text{SRA}}\bar{\mathbf{U}}^H\right). \tag{4.73}$$

Using the argument that the trace operator is invariant under cyclic permutations $\text{tr}(\mathbf{ABC}) = \text{tr}(\mathbf{CAB})$, it can be shown that:

$$\|\mathbf{P}_{\text{SRA}}\|^2_F = \sum_{i=1}^{N_r}\left(\frac{\bar{\sigma}_i}{\bar{\sigma}^2_i + N_r\sigma^2_{\widetilde{\mathbf{H}}}}\right)^2. \tag{4.74}$$

In order to quantify the maximum of $\|\mathbf{P}_{\text{SRA}}\|^2_F$, each component $\theta_i = \bar{\sigma}_i/\left(\bar{\sigma}^2_i + N_r\sigma^2_{\widetilde{\mathbf{H}}}\right)$ of the summation of (4.74) has to be upper bounded. Given that $\bar{\sigma}_i$ and $\sigma^2_{\widetilde{\mathbf{H}}}$ take positive values, it is clear to see that θ_i: i) is close to zero when $\bar{\sigma}_i \to 0$, ii) it attains the maximum value of $1/(1 + N_r\sigma^2_{\widetilde{\mathbf{H}}}) < 1$ when $\bar{\sigma}_i = N_r\sigma^2_{\widetilde{\mathbf{H}}}$, and iii) approaches to zero when $\bar{\sigma}_i \to +\infty$. Hence,

the instantaneous transmit power of $\mathbf{P}_{\text{SRA}}$ is always constrained by $\|\mathbf{P}_{\text{SRA}}\|_F^2 < N_r$, since $\bar{\sigma}_i/(\bar{\sigma}_i^2 + N_r\sigma_{\widetilde{\mathbf{H}}}^2) < 1$, for every positive value of $\bar{\sigma}_i$ and $\sigma_{\widetilde{\mathbf{H}}}^2$.

Again, following the same steps, the WCRA precoder is re-written as:

$$\mathbf{P}_{\text{WCRA}} = \bar{\mathbf{V}}\bar{\mathbf{\Sigma}}_{\text{WCRA}}\bar{\mathbf{U}}^H, \tag{4.75}$$

where,

$$\mathbf{\Sigma}_{\text{WCRA}} = \mathbf{diag}\left(\frac{\bar{\sigma}_1}{\bar{\sigma}_1^2 + \beta}, \ldots, \frac{\bar{\sigma}_{N_r}}{\bar{\sigma}_{N_r}^2 + \beta}\right). \tag{4.76}$$

The transmit power of WCRA is expressed as:

$$\|\mathbf{P}_{\text{WCRA}}\|_F^2 = \sum_{i=1}^{N_r}\left(\frac{\bar{\sigma}_i}{\bar{\sigma}_i^2 + \beta}\right)^2 < N_r. \tag{4.77}$$

Clearly, $\|\mathbf{P}_{\text{WCRA}}\|_F^2$ is also constrained because $\frac{\bar{\sigma}_i}{\bar{\sigma}_i^2+\beta} < 1$ for every $\{\bar{\sigma}_i, \beta\} > 0$.

A careful observation of (4.74) and (4.77) reveals that the channel uncertainty (in the form of $\sigma_{\widetilde{\mathbf{H}}}^2$ for SRA and in the form of β for WCRA) acts as a regularization parameter which adjusts the transmitted power. When the channel uncertainty approaches to zero, both SRA and WCRA reduce to ZF precoding and the transmitted power is maximized up to the transmitted power of ZF precoding. In contrast, when the channel uncertainty is increased, the level of power transmission imposed by the transmitter is decreased. This means that when the channel knowledge at the transmitter becomes more inaccurate, the transmitter reduces the transmitted power in order to avoid further degradation.

In the next few lines, the transmit power behaviour of ZF and MMSE precoding, when I-CSIT is available at the transmitter, is given without proof. Using a similar approach as before, it is shown that the transmit power of ZF precoding is:

$$\|\mathbf{P}_{\text{ZF}}\|_F^2 = \sum_{i=1}^{N_r}\frac{1}{\bar{\sigma}_i} \tag{4.78}$$

and the transmit power of MMSE precoding is:

$$\|\mathbf{P}_{\text{MMSE}}\|_F^2 = \sum_{i=1}^{N_r}\left(\frac{\bar{\sigma}_i}{\bar{\sigma}_i{}^2 + \mu}\right)^2. \tag{4.79}$$

From (4.78), it is clear that the power transmission of ZF precoding is increased to unacceptable high levels when the channel is in deep fade ($\bar{\sigma}_i^2 \rightarrow 0$). In contrast, for MMSE precoding, it can be shown that the power transmission is constrained by $\|\mathbf{P}_{\text{MMSE}}\|_F^2 < N_r$, because $\frac{\bar{\sigma}_i}{\bar{\sigma}_i^2+\mu} < 1$, $i = 1, \cdots, N_r$. A similar proof for the ZF and MMSE precoders is attained when the transmitter employs P-CSIT. In this case, the power transmission of the ZF and MMSE is given from (4.78) and (4.79) by replacing $\bar{\sigma}_i$ with σ_i, where σ_i^2 denote the eigenvalues of $\mathbf{H}^H\mathbf{H}$.

Finally, looking at (4.74), (4.77), and (4.79), it can be seen that there is a relation between the transmission power of different precoding methods. It is easy to demonstrate, using the structure of the power transmission of each precoder, that MMSE imposes higher instantaneous power transmission than: i) SRA when $\sigma_{\bar{\mathbf{H}}}^2 > \mu/N_r$, and ii) WCRA when $\beta > \mu$.

4.6.2 Average Transmit Power Analysis

Section 4.6.2 presents the analytical derivation of the average precoding power $\bar{P}_{\text{s}}$ of the methods employed in this thesis, when: i) I-CSIT is available and ii) $\mathbf{D} = \mathbf{I}_{N_r,N_r}$. Considering (4.74), (4.77), (4.78), and (4.79), it can be revealed that the instantaneous power of each precoder has the following structured form:

$$\|\mathbf{P}\|_F^2 = \sum_{i=1}^{N_r} \left(\frac{\bar{\sigma}_i}{\bar{\sigma_i}^2 + \vartheta} \right)^2 = \sum_{i=1}^{N_r} \left(\frac{\sqrt{\bar{\lambda}_i}}{\bar{\lambda}_i + \vartheta} \right)^2 . \tag{4.80}$$

Here, depending on the precoding method, ϑ takes the following values: i) $\vartheta = 0$ for ZF; ii) $\vartheta = \mu$ for MMSE; iii) $\vartheta = N_r\sigma_{\bar{\mathbf{H}}_{\text{SR}}}^2$ for SRA; and iv) $\vartheta = \beta$ for WCRA. In (4.80), for notational convenience, $\bar{\sigma_i}^2$ is substituted with $\bar{\lambda}_i$.

The analytical derivation of the average power of all precoding methods is given as:

$$\bar{P}_s = \text{E}\left[\|\mathbf{P}\|_F^2\right] = \text{E}_{\bar{\tau}}\left[\sum_{i=1}^{N_r} \left(\frac{\sqrt{\bar{\lambda}_i}}{\bar{\lambda}_i + \vartheta} \right)^2 \right], \tag{4.81}$$

where, $\bar{\tau} = \left[\bar{\lambda}_1, \ldots, \bar{\lambda}_{N_r}\right]^T$. It is not difficult to see that the random vector $\bar{\tau}$ contains the unordered eigenvalues of the Wishart matrix $\bar{\mathbf{W}} = \bar{\mathbf{H}}\bar{\mathbf{H}}^H \sim \mathcal{CW}(N_r, \mathbf{\Sigma}_{\bar{\mathbf{H}}})$, where, $\mathcal{CW}(N_r, \mathbf{\Sigma}_{\bar{\mathbf{H}}})$ denotes a central complex Wishart distribution as defined in [177, eq. 2.6]. Here, $\mathbf{\Sigma}_{\bar{\mathbf{H}}}$ is the covariance matrix of $\bar{\mathbf{H}}$. In the following, the expectation of (4.81) is derived when the wireless

channel is: i) uncorrelated and ii) transmit correlated ($\mathbf{R}_R = \mathbf{I}_{N_r,N_r}$).

4.6.2.1 Uncorrelated Channel

When the MIMO channel is uncorrelated, the covariance matrix of $\bar{\mathbf{H}}$ equals to $\mathbf{\Sigma}_{\bar{\mathbf{H}}} = \sigma^2_{\bar{\mathbf{H}}}\mathbf{I}$. Given that $\bar{\mathbf{H}} = \sqrt{\sigma^2_{\bar{\mathbf{H}}}}\mathbf{H}_w$, where, $\mathbf{H}_w \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$, the singular values $\bar{\lambda}_i$ of $\bar{\mathbf{H}}$ can be rewritten as $\bar{\lambda}_i = \sqrt{\sigma^2_{\bar{\mathbf{H}}}}\lambda_i$. Here, λ_i, $i = 1, \ldots, N_r$, represent the singular values of $\mathbf{H}_w$.

If the previous relation is taken into account by setting : i) $\bar{\tau} = \sqrt{\sigma^2_{\bar{\mathbf{H}}}}\tau$ and ii) $\phi = \vartheta/\sqrt{\sigma^2_{\bar{\mathbf{H}}}}$, after some manipulations, (4.81) is transformed to:

$$\bar{P}_s = \frac{1}{\sigma^2_{\bar{\mathbf{H}}}}\mathrm{E}_\tau\left[\sum_{i=1}^{N_r}\frac{\lambda_i}{(\lambda_i+\phi)^2}\right] = \frac{N_r}{\sigma^2_{\bar{\mathbf{H}}}}\mathrm{E}_\lambda\left[\frac{\lambda}{(\lambda+\phi)^2}\right]. \tag{4.82}$$

In (4.82), λ denotes the unordered eigenvalues of the complex central Wishart matrix $\mathbf{W} = \mathbf{H}_w\mathbf{H}_w^H$. The marginal PDF of λ is given in [177, eq. 2.23] as:

$$f_\lambda(x) = \frac{1}{N_r}\left[\sum_{k=0}^{N_r-1}\frac{k!}{(k+\Delta)!}\left[L_k^\Delta(x)\right]^2 x^\Delta e^{-x}\right]H_0(x), \tag{4.83}$$

where, $\Delta = N_t - N_r$, and $L_k^\Delta(x) = \sum_{i=0}^{k}(-1)^i\binom{k+\Delta}{k-i}\frac{x^i}{i!}$. Here, $L_k^\Delta(x)$ denotes the Laguerre polynomials. It is reminded that $H_0(x)$ is the Heaviside step function, for which it holds that $H_0(x) = 0$ for $x < 0$ and $H_0(x) = 1$ for $x \geq 0$.

Therefore, if the following expectation:

$$\mathrm{E}_\lambda\left[\frac{\lambda}{(\lambda+\phi)^2}\right] = \int_0^{+\infty}\frac{x}{(x+\phi)^2}f_\lambda(x)\,\mathrm{d}x, \tag{4.84}$$

is evaluated using (4.83), the average precoding power of (4.82) becomes:

$$\begin{aligned}\bar{P}_s = \frac{1}{\sigma^2_{\bar{\mathbf{H}}}}\sum_{k=0}^{N_r-1}\frac{k!}{(k+\Delta)!}\sum_{l=0}^{k}\sum_{m=0}^{k}\frac{(-1)^{l+m}}{l!m!}\binom{k+\Delta}{k-l}\\ \times\binom{k+\Delta}{k-m}(n+1)!\phi^n e^\phi\left[\Gamma(-n,\phi)-\phi\Gamma(-n-1,\phi)\right].\end{aligned} \tag{4.85}$$

In (4.85), $n = l + m + \Delta$ and $\Gamma(\cdot\,,\cdot)$ denotes the incomplete gamma function defined in [170, p. 899]. The evaluation of the integral of (4.84) is given in Appendix A.

4.6.2.2 Transmit Correlated Channel

The elaboration of (4.81) gives that:

$$\bar{P}_s = \mathrm{E}_{\bar{\tau}}\left[\sum_{i=1}^{N_r}\left(\frac{\sqrt{\bar{\lambda}_i}}{\bar{\lambda}_i+\vartheta}\right)^2\right] = N_r \mathrm{E}_{\bar{\lambda}}\left[\frac{\bar{\lambda}}{(\bar{\lambda}+\vartheta)^2}\right], \tag{4.86}$$

where, $\bar{\lambda}$ is the unordered eigenvalues of the complex central Wishart matrix $\bar{\mathbf{W}} = \bar{\mathbf{H}}\bar{\mathbf{H}}^H$. Clearly, the evaluation of (4.86) requires the marginal PDF of $\bar{\lambda}$, which is derived below.

The presence of only transmit space correlation in the wireless channel determines that $\bar{\mathbf{H}} = \mathbf{H}_w\left(\bar{\mathbf{R}}_T^{1/2}\right)^T$, where, $\bar{\mathbf{R}}_T^{1/2} = \sqrt{\sigma^2_{\bar{\mathbf{H}}}}\left(\mathbf{R}_T^{1/2}\right)^T$. Thus, $\bar{\mathbf{H}}$ is distributed according to the following complex Gaussian distribution $\mathbf{H} \sim \mathcal{CN}(\mathbf{0}, \bar{\mathbf{R}}_T \otimes \mathbf{I}_{N_r,N_r})$. In this case, the complex central Wishart matrix $\bar{\mathbf{W}}$ is distributed as $\mathcal{CW}(N_R, \bar{\mathbf{R}}_T)$. For this reason, as it shown in [178], the marginal PDF of $\bar{\lambda}$ equals to:

$$f_{\bar{\lambda}}(x) = K\left[\sum_{i=1}^{N_r}\sum_{j=1}^{N_r} x^{j-1}\mathcal{D}(i,j)\left[\alpha_{\varrho+i}^{\varrho-1}e^{\frac{-x}{\alpha_{\varrho+i}}} - \sum_{l=1}^{\varrho}\sum_{k=1}^{\varrho}\left[\mathbf{\Psi}^{-1}\right]_{k,l}\alpha_{\varrho+i}^{k-1}\alpha_l^{\varrho-1}e^{-\frac{x}{\alpha_l}}\right]\right], \tag{4.87}$$

where, $\varrho = N_t - N_r$ and

$$K = \frac{\det(\mathbf{\Psi})}{N_r\left[\prod_{k<l}^{N_t}(\alpha_l - \alpha_k)\right]\prod_{l=1}^{N_r-1} l!}.$$

Here, $\mathbf{\Psi}$ is the following $\varrho \times \varrho$ Vandermonde matrix:

$$\Psi = \begin{bmatrix} 1 & \alpha_1 & \cdots & \alpha_1^{\varrho-1} \\ \vdots & \ddots & \ddots & \vdots \\ 1 & \alpha_\varrho & \cdots & \alpha_\varrho^{\varrho-1} \end{bmatrix}.$$

In addition, $\alpha_1, \ldots, \alpha_\varrho$, represent the ϱ distinct eigenvalues of $\bar{\mathbf{R}}_T$. The (i,j) cofactor of the $N_r \times N_r$ matrix $\mathbf{C}$ whose (l,k) element equals to:

$$[\mathbf{C}]_{l,k} = (k-1)!\left(\alpha_{\varrho+l}^{\varrho+k} - \sum_{p=1}^{\varrho}\sum_{q=1}^{\varrho}\left[\mathbf{\Psi}^{-1}\right]_{p,q}\alpha_{\varrho+l}^{p-1}\alpha_q^{\varrho+k-1}\right)$$

is denoted as $\mathcal{D}(i,j)$.

At this point, given that the marginal PDF of $\bar{\lambda}$ is available in (4.87), the following expectation can be evaluated:

$$\mathrm{E}_{\bar{\lambda}}\left[\frac{\bar{\lambda}}{(\bar{\lambda}+\vartheta)^2}\right]=\int_0^{+\infty}\frac{x}{(x+\vartheta)^2}f_{\bar{\lambda}}(x)\mathrm{d}x \tag{4.88}$$

and plugged in (4.86). The evaluation of the integral of (4.88) is given in Appendix B. Hence, it can be proven that:

$$\bar{P}_s=KN_r\sum_{i=1}^{N_r}\sum_{j=1}^{N_r}\mathcal{D}(i,j)\left[\alpha_{\varrho+i}^{\varrho-1}g_j\left(\frac{1}{\alpha_{\varrho+i}}\right)-\sum_{l=1}^{\varrho}\sum_{k=1}^{\varrho}\left[\boldsymbol{\Psi}^{-1}\right]_{k,l}\alpha_{\varrho+i}^{k-1}\alpha_l^{\varrho-1}g_j\left(\frac{1}{\alpha_l}\right)\right], \tag{4.89}$$

where,

$$g_j\left(y\right)=\vartheta^{j-1}e^{\vartheta y}j!\left[\Gamma\left(1-j,y\vartheta\right)-y\vartheta\Gamma\left(-j,y\vartheta\right)\right] \tag{4.90}$$

and $\Gamma(\cdot)$ denotes the gamma function defined in [170, p. 892].

4.7 Results and Discussion

Analytical and Monte Carlo simulation results which evaluate the performance of R-SM and MSR-SM are provided in this section. Initially, it is considered a scenario where the transmitter possesses P-CSIT and the precoding method is ZF. In this case, the wireless channel is flat and follows a Rayleigh distribution. The tightness of the derived bounds of Sections 4.3.1 and 4.3.2 is verified. In addition, the performance of R-SM and MSR-SM is compared against the performance of the corresponding conventional MIMO setup. In this way, it is shown which scheme performs better.

The assumption of P-CSIT is too idealistic. In a real scenario CSIT is subject to imperfections. Thus, in order to study the performance of R-SM and MSR-SM when I-CSIT is available, the channel model of Section 4.5 and the precoding methods of Sections 4.5.3 and 4.5.3. are deployed.

4.7.1 Validation of the Theoretical Average Bit Error Probability

In this section, simulation results that confirm the tightness of the bounds of the ABEP of Sections 4.3.1 and 4.3.2 are provided. For the bounds of R-SM presented in Section 4.3.1, the considered system setups are: i) a 4×2 ($N_t=4$ and $N_r=2$) with 4 bpcu spectral efficiency; and ii) a 8×4 ($N_t=8$ and $N_r=$) system configuration which transmits 4 bpcu. Furthermore,

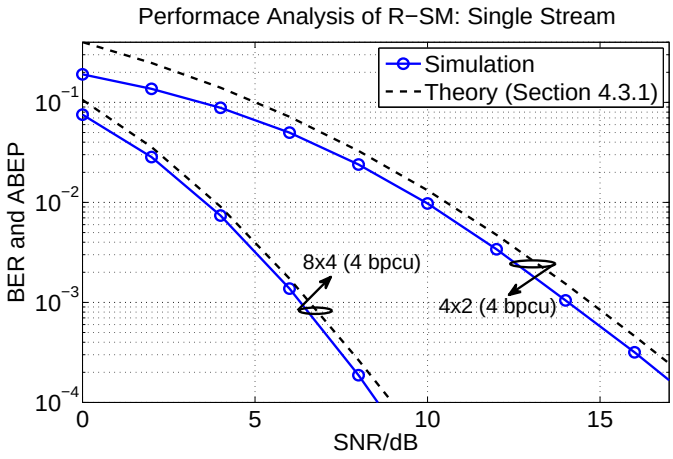


Figure 4.3: *Performance analysis of R-SM (single stream) when ZF with P-CSIT is employed at the transmitter, using: i) simulation results and ii) the bounds of Section 4.3.2.*

the tightness of the bounds of Section 4.3.2 of MSR-SM is presented for the following system configurations: i) a 4×4 ($N_t = 4$ and $N_r = 4$) system setup with one stream ($N_s = 1$) and 4 bpcu spectral efficiency; ii) a 4×4 ($N_t = 4$ and $N_r = 4$) system setup with two streams ($N_s = 2$) and 6 bpcu; and iii) a 5×3 ($N_t = 5$ and $N_r = 3$) system setup with two streams ($N_s = 2$) and 6 bpcu spectral efficiency. Note that all system setups deploy a QAM constellation with order 4. Furthermore, for all simulations presented in this chapter, it is assumed that $\mathrm{E}_{\mathbf{x}}[\mathbf{x}] = 1$.

As shown in Figs. 4.3 and 4.4, the proposed upper bounds of Section 4.3.1 and 4.3.2, respectively, are tight and capture the nature of the behavior of the performance of R-SM and MSR-SM in high SNR. In low SNR, there is a deviation which is a well known effect of the union bound technique [169]. Furthermore, the analysis of the diversity orders of R-SM and MSR-SM presented in Section 4.4 are confirmed from Figs. 4.3 and 4.4, respectively. In both figures, the slope of the BER curves, which are obtained via simulation results, is in perfect agreement with the theoretical analysis of Section 4.4, as the SNR approaches to infinity.

4.7.2 BER Performance of Receive-Spatial Modulation

For the purpose of comparison, Fig. 4.5 depicts the BER performance of R-SM and the performance of the corresponding conventional MIMO scheme, for the same system setups. Here, the conventional MIMO scheme deploys ZF precoding with P-CSIT. Its operating principle determines that N_r parallel and non interfering streams are established from the transmitter to

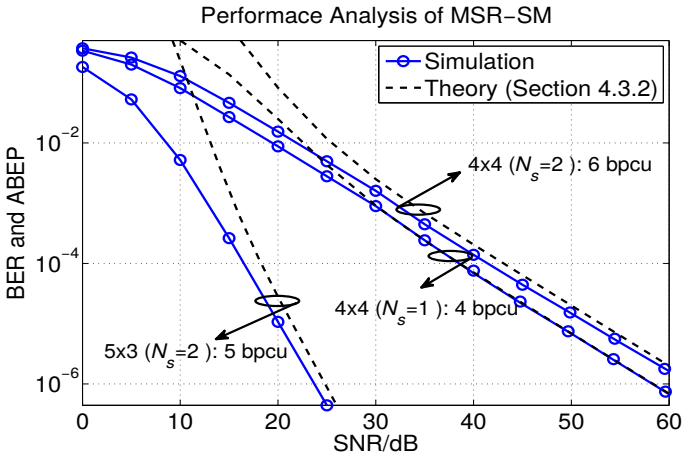


Figure 4.4: *Performance analysis of MSR-SM (single stream and multiple streams) when ZF with P-CSIT is employed at the transmitter, using: i) simulation results and ii) the bounds of Section 4.3.1.*

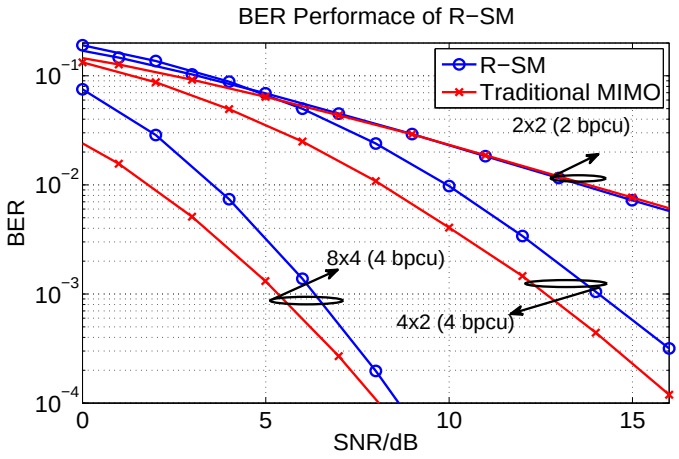


Figure 4.5: *BER performance of R-SM versus conventional MIMO.*

the receiver. Furthermore, detection at the receiver is conducted using the ML criterion. Note that all conventional symbols, both for R-SM and conventional MIMO are drawn from a M-ary QAM or BPSK constellation diagrams. The value of M is selected such that the desired spectral efficiency is guaranteed. As shown in Fig. 4.5, both schemes have the same performance for a 2×2 and 2 bpcu system setup. Whereas, for a system setup 4×2 (4 bpcu) and a system setup 8×4 (4 bpcu), conventional MIMO outperforms R-SM, due to its higher multiplexing gain. However, from Fig. 4.5, it can be indirectly concluded that both systems achieve the same diversity order. Indeed, it is verified theoretically that their diversity equal to $L = N_t - N_r + 1$. Note that the derivation of the diversity order for R-SM is undertaken in Section 4.4, while the diversity order of conventional MIMO is derived in [70].

4.7.3 BER Performance of Multi-Stream Receive-Spatial Modulation

The BER performance of MSR-SM is studied in this section. Initially, the effect of the number of spatially modulated parallel data streams N_s in the performance of MSR-SM is explored. Fig. 4.6, presents the BER performance of a 8×4 ($N_t = 8$ and $N_r = 4$) system for different values of $N_s = \{1, 2, 3\}$. For all system setups, the spectral efficiency is set to 8 bpcu by selecting the appropriate order for the deployed QAM constellation. As shown in Fig. 4.6, as the number of parallel spatially modulated data streams N_s is increased, MSR-SM performs better. In most cases, for the same spectral efficiency, as the number of parallel stream is increased, the order of the deployed constellation is decreased. Therefore, the performance becomes better. For example, in Fig. 4.6, when $N_s = 1$ (one stream), MSR-SM requires 64-QAM, while when $N_s = 2$ (two streams) the order of the deployed QAM constellation is 8, and when $N_s = 3$ (three stream) MSR-SM uses 4-QAM. However this is not a general rule. There are some special cases where different numbers of N_s result in the same spectral efficiency using the same order of constellation. In this case, MSR-SM achieves better BER performance when N_s is smaller. An example is presented in Fig. 4.7. More details are given below.

Furthermore, Fig. 4.6 presents the comparison of MSR-SM with the corresponding conventional MIMO scheme. The definition of the benchmark system is given in Section 4.7.2. As shown in Fig. 4.6, the benchmark system outperforms MSR-SM when N_s equals to one and two, due to its higher multiplexing gain. In fact, provided that the benchmark system establishes four parallel data streams from the transmitter to the receiver, the required order of the deployed QAM constellation is only four, for 8 bpcu. In contrast, when $N_s = 3$, MSR-SM uses 4-QAM, which is the same order as the benchmark system, in order to transmit 8 bpcu. For this reason, in high SNR, MSR-SM has a better performance than the benchmark system. For a BER $\approx 10^{-6}$, MSR-SM outperforms conventional MIMO for about 1.1 dB. Hence, in the following, the focus is given on the comparison of MSR-SM with conventional MIMO when the value N_s is high.

Fig. 4.7 presents the BER performance of MSR-SM for the following system setups: i) 8×5 with $N_s = 4$ and 10 bpcu; ii) 8×6 with $N_s = 5$ and 12 bpcu; ii) 8×7 with $N_s = 5$ and 14 bpcu; and iv) 8×7 with $N_s = 6$ and 14 bpcu. For comparison, Fig. 4.7 gives the performance of the corresponding conventional MIMO scheme, as defined in Section 4.7.2. For all schemes in Fig. 4.7, the order of the utilized QAM constellation is 4. Fig. 4.7 shows that in low SNR, the benchmark system offers a slightly better BER than MSR-SM. However, as the SNR is

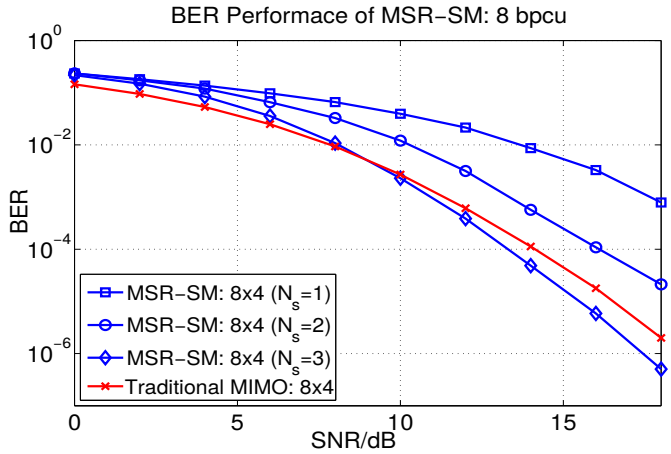


Figure 4.6: *BER performance of MSR-SM for a* 8×4 *system and* $N_s = \{1, 2, 3\}$ *(number of parallel spatially modulated data streams).*

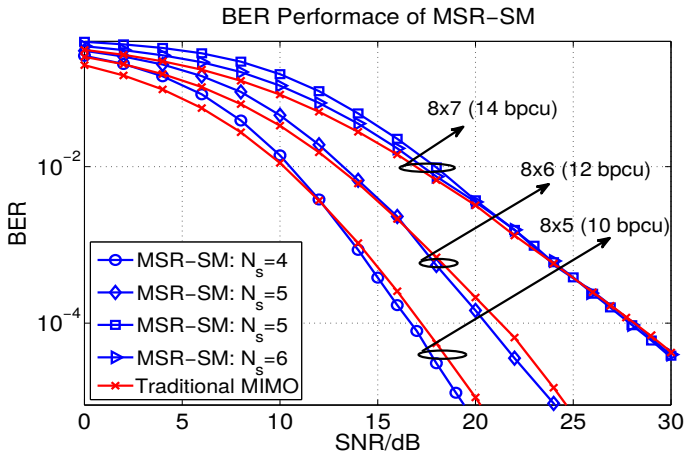


Figure 4.7: *BER performance of a MSR-SM versus conventional MIMO.*

increased in values of practical interest, there is a crossing point above which the benchmark system is outperformed. The crossing point for the system setups of 8×5 with $N_s = 4$, 8×6 with $N_s = 5$, and 8×7 with $N_s = 6$ occurs at 12.5, 16.5, and 27.5 dB, respectively. Note that even though there is a crossing point, both MSR-SM and the benchmark system achieve the same diversity order ($L = N_t - N_r + 1$). Indeed, if the SNR is increased up to infinity, there is a crossing point after which the BER curves of both systems are parallel (same slope). It is reminded that the diversity order for MSR-SM is obtained in Section 4.4 and the diversity order for the benchmark system is given in [70].

Fig. 4.7 also shows that for a system setup of 8×7, with N_s equals to 5 or 6, the same spectral efficiency of 14 bpcu is achieved using the same QAM order. In this case, the performance of

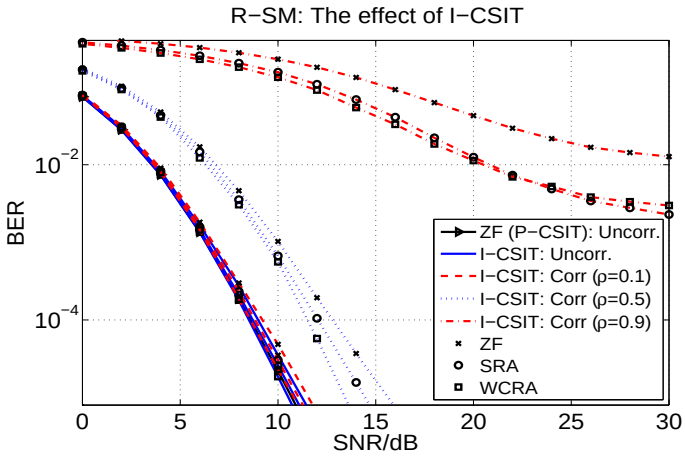


Figure 4.8: *BER performance of R-SM (8×4 system) when I-CSIT is available.*

MSR-SM with respect to N_s depends on the value of SNR. Note that this is a very special case, where different values of N_s result to the same spectral efficiency, when the QAM order is the same. For values of SNR below 9.2 dB and above 27.5 dB, MSR-SM performs better when $N_s = 6$. In contrast, for values of SNR between 9.2 and 27.5 dB, setting $N_s = 6$ is preferable. This phenomenon is attributed to the fact, that even though the values of N_s are different and the QAM order is the same, the resultant spectral efficiency is same. This is due the different number of bits offered by different values of N_s.

4.7.4 BER performance: The Effect of I-CSIT

In this section, using simulation results, the performance of R-SM and MSR-SM is studied when the transmitter deploys I-CSIT. The system configuration assumes $N_t = 8$ transmit antennas and $N_r = 4$ receive antennas. For R-SM, it is assumed that $N_s = 1$, while for MSR-SM it is assumed that $N_s = 3$. Furthermore, for both R-SM and MSR-SM, the utilized constellation is 4-QAM. Therefore, the spectral efficiency for R-SM is 4 bpcu, while the spectral efficiency for MSR-SM is 8 bpcu. The MIMO channel $\mathbf{H}$ is generated as the sum of two Zero Mean Circular Symmetric Complex Gaussian (ZMCSCG) matrices ($\overline{\mathbf{H}}$ and $\widetilde{\mathbf{H}}$) whose elements follow i.i.d. ZMCSCG distributions ($\overline{h}_{i,j} \sim \mathcal{CN}(0, 0.99)$ and $\tilde{h}_{i,j} \sim \mathcal{CN}(0, 0.01)$) such that $h_{i,j} \sim \mathcal{CN}(0, 1)$. The entries of $\overline{\mathbf{H}}$ are refreshed every 100 symbol periods, whereas the entries of $\widetilde{\mathbf{H}}$ are refreshed every symbol period. In addition, the entries of the spatial correlation matrices $\mathbf{R}_R$ and $\mathbf{R}_T$ are generated using an exponential model with $\rho_t = \rho_r = \rho$, where ρ takes values from the set $\{0.1, 0.5, 0.9\}$.

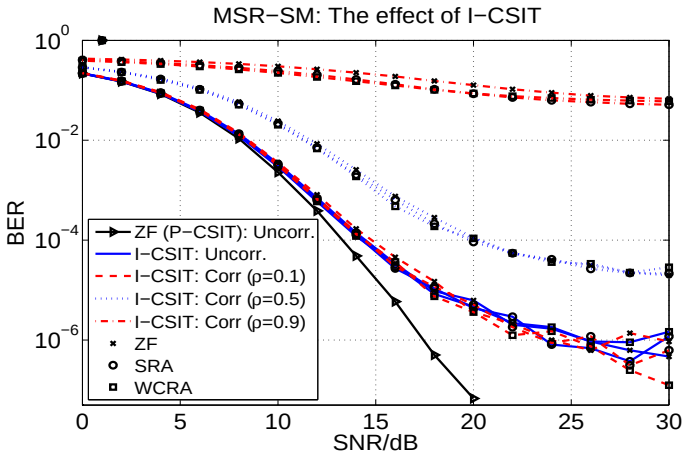


Figure 4.9: *BER performance of MSR-SM (8×4 system with $N_s = 3$) when I-CSIT is available.*

The BER performance of R-SM and MSR-SM, for different precoding methods (ZF, SRA, and WCRA), I-CSIT, and for the cases of uncorrelated and correlated channel, is presented in Figs. 4.8 and 4.9. Especially, for the correlated case, spatially correlated channels with low correlation ($\rho = 0.1$), with intermediate correlation ($\rho = 0.5$) and with high correlation ($\rho = 0.9$) are assumed. In addition, for comparison purposes, Figs. 4.8 and 4.9 give the BER performance of the ZF precoder under P-CSIT and I-CSIT. In order to conduct a fair comparison between the different precoding designs, the wireless channel follows the same statistical characteristics for all cases.

The observation of Figs. 4.8 and 4.9 shows that under an uncorrelated channel or low correlation ($\rho = 0.1$), R-SM and MSR-SM with I-CSIT, and for all precoding designs, have almost the same performance as ZF precoding with P-CSIT in low SNR. Therefore, in such a scenario, the performance of R-SM and MSR-SM does not deviate from the ideal channel conditions. Furthermore, as shown in Fig. 4.9, for the depicted values of SNRs, the performance of R-SM with I-CSIT follows the ideal case. In contrast, in high SNR, the performance of MSR-SM with I-CSIT is outperformed from the corresponding performance of ZF precoding with I-CSIT. In high SNR, the main degradation factor under the presence of I-CSIT is the residual ICI.

Fig. 4.9 demonstrates that the main effect of I-CSIT on the performance of MSR-SM is a bottleneck, where higher values of SNR do not affect its performance substantially. This phenomenon occurs in all scenarios of channel correlation. In contrast, Fig. 4.8 shows that the performance of R-SM faces a bottleneck, for the depicted SNRs, only in high correlation ($\rho = 0.9$). However, due to the fact that the presence of I-CSIT results in a noise limited regime (residual ICI

at the receiver), a bottleneck is expected in higher values of SNR.

As shown in Figs. 4.8 and 4.9, the way that correlation affects the performance of R-SM and MSR-SM can be divided in two cases. In the low correlation case ($\rho = 0.1$), it can be seen that there is no performance difference between the uncorrelated and correlated channel. In contrast, as correlation increases, the performance of R-SM and MSR-SM deteriorates. For the intermediate level of correlation ($\rho = 0.5$) and BER=10^{-3}, there is a SNR penalty of about 4 dB for R-SM (Fig. 4.8) and 4.5 dB for MSR-SM (Fig. 4.9). For the high level of correlation ($\rho = 0.9$), both R-SM and MSR-SM are not able to achieve a BER less than 10^{-3} in the depicted values of SNR.

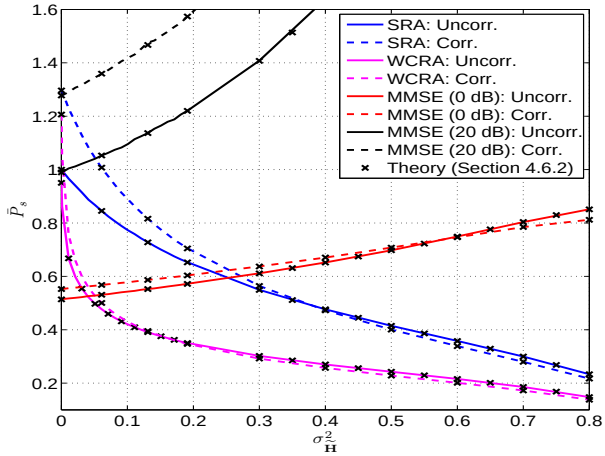
4.7.5 Power Transmission and Residual MSE at the Receiver when I-CSIT is Available

In Section 4.6, it is demonstrated that the unnormalized MMSE, SRA, and WCRA precoders result in a constrained power transmission, while the ZF precoder does not. Hence, when the MMSE, SRA, and WCRA precoding designs are employed in a real system implementation, it is possible to avoid the use of the precoding matrix $\mathbf{D}$ at the transmitter. In this case, given the assumption of $\mathbf{HP} \approx \mathbf{I}_{N_r,N_r}$, there is no need for the receiver to have the knowledge of d_i, $i = 1, \dots N_r$. Thus, a non-coherent detection algorithm is able to be deployed at the receiver.

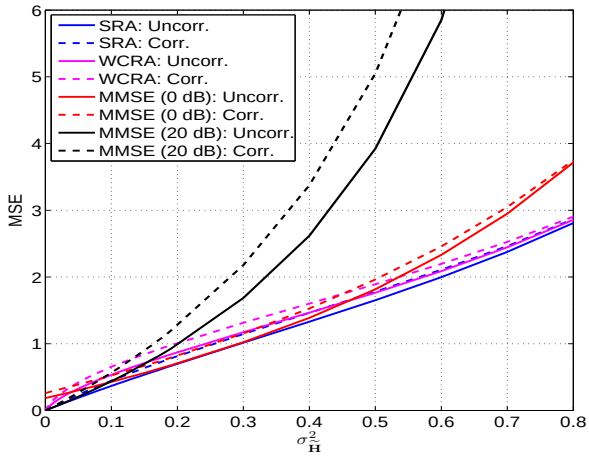
In Fig. 4.10, the average transmitted power and the residual Mean Square Error (MSE) at the receiver are evaluated with respect to the variance of channel uncertainty $\sigma^2_{\widetilde{\mathbf{H}}}$. This is done using simulation results and the theoretical framework of Section 4.6.2. In addition, this is done for all precoding designs (MMSE, SRA, and WCRA) when the normalization matrix $\mathbf{D}$ is not applied. Thus, the transmitted power is defined as $P_{\mathrm{s}} = \mathrm{E}_{\mathbf{x}}\left[\|\mathbf{Px}\|_2^2\right]$.

As shown in Fig. 4.10(a), when $\sigma^2_{\widetilde{\mathbf{H}}}$ is increased, the transmitted power of SRA and WCRA precoding methods is decreased. This means that, as the CSIT becomes more inaccurate, the power transmission is reduced in order to avoid further degradation (inaccurate beamforming). In contrast, when MMSE precoding is employed and for all the depicted values of SNR $= \{0, 20\}$ dB, power transmission is increased as the variance of channel uncertainty is increased. This results to inaccurate beamforming.

Furthermore, Fig. 4.10(a) shows that correlation affects the transmitted power of the new precoding design in a composite way. When $\sigma^2_{\widetilde{\mathbf{H}}} \to 0$, higher correlation (higher values of ρ)



(a) Average transmitted power.



(b) Residual MSE at the receiver.

Figure 4.10: *Effect of I-CSIT to a* 8×4 *system.*

results in higher power transmission. In contrast, when $\sigma^2_{\tilde{\mathbf{H}}}$ is increased, higher correlation results in lower power transmission. In contrast to the new designs, the existence of channel correlation, when MMSE precoding is used, imposes higher power transmission. As a final remark, it is noted that all of the theoretical curves perfectly match with the simulated curves. In addition, all the claims of Section 4.6 concerning the relation between the proposed precoding designs and MMSE precoding are confirmed.

Fig. 4.10(b) demonstrates the average residual MSE at the receiver, defined as:

$$\text{MSE} = \|\mathbf{HP} - \mathbf{I}_{N_r,N_r}\|_F^2 \tag{4.91}$$

where, $\mathbf{I}_{N_r,N_r}$ is the ideal ZF impulse response. Fig. 4.10(b) quantifies the residual ICI caused by the inaccurate beamforming, versus the variance of channel uncertainty. As it can be seen, for values of $\sigma^2_{\widetilde{\mathbf{H}}}$ close to zero (almost no uncertainty), MMSE results to lower MSE. It is well known that the MMSE precoder is the optimum linear precoder in terms of MSE. However, as the channel becomes more inaccurate (higher values of $\sigma^2_{\widetilde{\mathbf{H}}}$), the new precoding designs offer significantly lower residual MSE than MMSE (for all SNR values). The same figure depicts the performance of the studied precoding designs when the channel is correlated. Clearly, higher correlation results in higher MSE. Finally, it can be seen that as the channel uncertainty becomes high, the effect of correlation to SRA and WCRA is insignificant compared to the uncorrelated case.

4.7.6 Energy Efficiency, Complexity, and Data Rate Trade-off Analysis Between Receive Space Modulation and Conventional MIMO

In this section, a trade-off analysis between the EE, complexity, and data rate of R-SM, MSR-SM, and the corresponding conventional MIMO benchmark system is provided. Note that the description of the benchmark system is provide in Section 4.7.2.

In order to quantify the EE benefits of the studied schemes, the metric of Relative Average Energy Reduction (RAER) is deployed [179]. In particular, the metric of RAER is defined as follows:

$$\text{RAER}[\%] = \left[1 - 10^{-\frac{\Delta_{\text{SNR}}}{10}}\right] \times 100\%, \tag{4.92}$$

where, Δ_{SNR} denotes the SNR difference (in dB) between R-SM or MSR-SM and the conventional MIMO benchmark system, for a given BER. It is worth mentioning that only the energy consumption for the RF power transmission is considered. Both architectures, in fact, have almost the same circuits energy consumption, since they employ the same number of RF front-ends.

Similarly, the complexity benefits of the studied schemes are quantified by the metric of Relative Complexity Reduction (RCR), defined as:

$$\text{RCR}[\%] = \frac{\mathcal{C}_t^{\text{MIMO}} - \mathcal{C}_t^{\text{Space Modulation}}}{\mathcal{C}_t^{\text{MIMO}}} \times 100\%, \tag{4.93}$$

where, $\mathcal{C}_t^{\text{MIMO}}$ and $\mathcal{C}_t^{\text{Space Modulation}}$, is the computational complexity (in real additions and

Configuration	k_{system} (bpsp)	Target BER	RAER [%]	RCR [%]
$N_t = 2$, $N_r = 2$, $N_s = 1$	2	10^{-2}	4.5%	56%
$N_t = 8$, $N_r = 4$, $N_s = 1$	4	10^{-3}	−25%	79%
$N_t = 8$, $N_r = 4$, $N_s = 3$	8	10^{-4}	18%	26%
$N_t = 8$, $N_r = 4$, $N_s = 2$	8	10^{-4}	−58%	53%

Table 4.1: *Energy Efficiency, Complexity, and Data Rate Trade-off Analysis of receive space modulation with respect to the conventional MIMO benchmark system.*

multiplications) of computing the transmitted signal $\mathbf{s} = \mathbf{PDx}$, when the MIMO benchmark system and R-SM or MSR-SM are deployed, respectively. By taking into account the sparse structure of the transmitted signal vector $\mathbf{x}$ of R-SM and MSR-SM in (4.4) and (4.6), respectively, it can be shown that the transmit computational complexity of R-SM is:

$$\mathcal{C}_t^{\text{R}-\text{SM}} = 6N_t + 2 \tag{4.94}$$

and the transmit computational complexity of MSR-SM is:

$$\mathcal{C}_t^{\text{MSR}-\text{SM}} = N_t(8N_s - 2) + 2N_s. \tag{4.95}$$

Finally, the transmit computational complexity of the benchmark system is given as:

$$\mathcal{C}_t^{\text{MIMO}} = N_t(8N_r - 2) + 2N_r. \tag{4.96}$$

As shown in Table 4.1, for the same data rate, R-SM and MSR-SM have a lower computational complexity, than the MIMO benchmark system. As regard the metric of RAER, as it can be seen in Table 4.1, in some system setups, the corresponding MIMO benchmark system is more energy efficient than R-SM and MSR-SM. However, for $N_t = 2$ and $N_r = 2$, R-SM is more energy efficient than the benchmark system. Also, for $N_t = 8$, $N_r = 4$, and $N_s = 3$, MSR-SM offers a higher EE than the benchmark system.

4.8 Summary

In this chapter, using theoretical and simulation results, the performance of R-SM and MSR-SM was evaluated. In more detail, a theoretical upper bound of the ABEP of R-SM and MSR-SM were provided, when: i) the wireless channel is uncorrelated and follows the Rayleigh distribution; and ii) the precoder is designed using the ZF principle with P-CSIT. In addition,

the diversity order and coding gain of both schemes were derived. A comparison was given for R-SM and MSR-SM against the corresponding conventional MIMO scheme. From these comparisons, it was concluded when they perform better. Furthermore, with the assumption of I-CSIT, two linear precoders were proposed. The instantaneous and average precoding power for all precoding methods considered in this thesis were analyzed theoretically.

Chapter 5
On the Performance of Multi-Stream Receive-Spatial Modulation in the MIMO Broadcast Channel

5.1 Introduction

In this chapter, the incorporation of MSR-SM for application to the MIMO broadcast channel is introduced and its BER performance is mathematically studied. More specifically, based on the union bound technique, an accurate mathematical framework for its performance evaluation is proposed and discussed. From this framework, it is proved that MSR-SM provides the same diversity order as the conventional MIMO broadcast channel, while offering a better coding gain in the high SNR regime. Also, this performance gain is achieved with a reduction of the complexity of the transmitter. Numerical simulations are shown in order to substantiate the gain predicted by the analysis. As a result, MSR-SM is shown to be a promising transmission scheme for the MIMO broadcast channel.

5.2 System Model

An uncoded Multi-User Multiple-Input Multiple-Output (MU-MIMO) system that comprises a multi-antenna BS and N_u remotely distributed multi-antenna users is considered. Such a system is shown in Fig. 5.1. The BS is equipped with N_t antennas and each user possesses N_r antennas. Since the transmitter is a BS, it is realistic to consider the assumption that $N_t \geq N_u N_r$. In addition, the wireless channel between the BS and every user is assumed to be frequency flat and quasi-static. Finally, P-CSIT is considered, which can be obtained by using either the channel reciprocity or fast and error free links from the users[1].

[1] In real systems, CSIT is subjected to imperfections. However, the study of the effect of I-CSIT is outside the scope of this chapter.

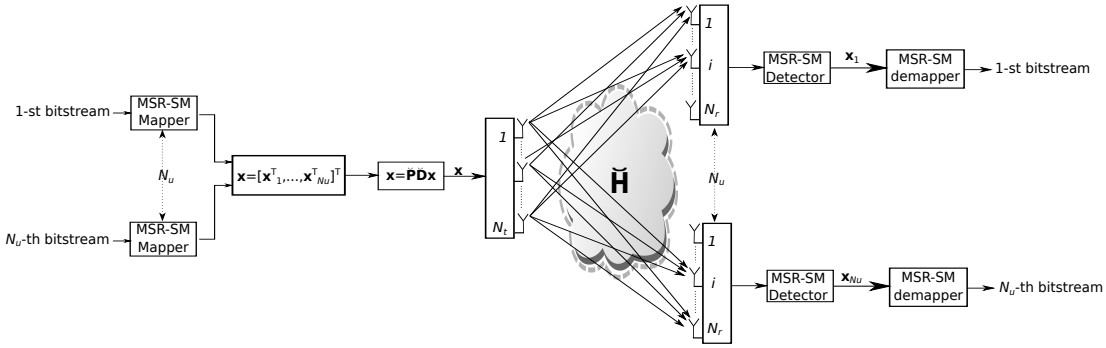


Figure 5.1: *The system model of MU MSR-SM.*

Provided P-CSIT is available, the transmitter is able to use linear precoding. By interpreting the N_t transmit antennas and the $B = N_u N_r$ receive antennas as a $N_t \times B$ MIMO system, the baseband equation of such a system is expressed as:

$$\mathbf{y} = \breve{\mathbf{H}}\breve{\mathbf{P}}\breve{\mathbf{D}}\mathbf{x} + \mathbf{w}, \tag{5.1}$$

in a matrix form. In (5.1), $\mathbf{y}$ is a $N_u N_r \times 1$ vector which is written as:

$$\mathbf{y} = \left[\mathbf{y}_1^T, \ldots, \mathbf{y}_{N_u}^T\right]^T, \tag{5.2}$$

where $\mathbf{y}_i$, $i = 1, \ldots, N_u$, denotes the $N_r \times 1$ received signal vector at the i-th user. The $N_u N_r \times N_t$ matrix:

$$\breve{\mathbf{H}} = \left[\breve{\mathbf{H}}_1^H, \ldots, \breve{\mathbf{H}}_{N_u}^H\right]^H \tag{5.3}$$

denotes the wireless channel from the transmitter to all receive antennas. Furthermore, the sub-matrix $\breve{\mathbf{H}}_i$, $i = 1, \ldots, N_u$, denotes the wireless channel from the transmitter to the i-th user. Due to the spatial distribution of the users inside the geographical area of a cell, each user experiences a different large scale effect. However, given that the receive antennas of each user are collocated, the large scale effect between the transmitter and each receive antenna of a certain user is the same. Therefore, it is assumed that $\breve{\mathbf{H}}_i$, $i = 1, \ldots, N_u$, has the following distribution: $\breve{\mathbf{H}}_i \sim \mathcal{CN}(\mathbf{0}, \xi_i \mathbf{I})$. The values of ξ_i close to zero represent a poor channel condition, while those of ξ_i close to one indicate a strong channel condition. It is assumed that there is no channel correlation due to rich scattering. Thus, (5.3) is expressed as:

$$\breve{\mathbf{H}} = \mathbf{\Xi}^{\frac{1}{2}}\mathbf{H}, \tag{5.4}$$

where,

$$\mathbf{\Xi} = \mathbf{diag}\left(\xi_1\mathbf{I}, \ldots, \xi_{N_u}\mathbf{I}\right), \tag{5.5}$$

represents the large scale effect. In addition, $\mathbf{H}$ is defined as:

$$\mathbf{H} = \left[\mathbf{H}_1^H, \ldots, \mathbf{H}_{N_u}^H\right]^H, \tag{5.6}$$

where, $\mathbf{H} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ and $\mathbf{H}_i$, $i = 1, \ldots, N_u$, represents the small scale fading of the i-th user. The $N_t \times N_u N_r$ precoding matrix can be formulated as:

$$\breve{\mathbf{P}} = \left[\breve{\mathbf{P}}_1, \ldots, \breve{\mathbf{P}}_{N_u}\right], \tag{5.7}$$

where, $\breve{\mathbf{P}}_i$, $i = 1, \ldots, N_u$, corresponds to the precoding matrix of the i-th user. In order to ensure that the transmitted power is not amplified by the precoder $\breve{\mathbf{P}}$, a $N_u N_r \times N_u N_r$ diagonal normalization matrix $\breve{\mathbf{D}} = \mathbf{diag}\left(\breve{d}_1, \ldots, \breve{d}_{N_u N_r}\right)$ is used. Every element $\breve{d}_i$, $i = 1, \ldots, N_u N_r$, of $\breve{\mathbf{D}}$ is expressed as:

$$\breve{d}_i = \sqrt{\frac{1}{\|\breve{\mathbf{p}}_i\|_2^2}}, \tag{5.8}$$

where, $\breve{\mathbf{p}}_i$ is the i-th column of $\breve{\mathbf{P}}$. Thus, every column of the normalized precoding matrix

$$\breve{\mathbf{P}}_{\text{norm}} = \breve{\mathbf{P}}\breve{\mathbf{D}} \tag{5.9}$$

has unity power. Thus, the normalization matrix $\breve{\mathbf{D}}$ can be expressed in the following block diagonal matrix form:

$$\breve{\mathbf{D}} = \mathbf{diag}\left(\breve{\mathbf{D}}_1, \ldots, \breve{\mathbf{D}}_{N_u}\right). \tag{5.10}$$

In (5.10), $\breve{\mathbf{D}}_i$, $i = 1, \ldots, N_u$, is the $N_r \times N_r$ diagonal normalization matrix of the corresponding precoding matrix $\breve{\mathbf{P}}_i$. The collective signal vector at the transmitter is denoted as:

$$\mathbf{x} = \left[\mathbf{x}_1^T, \ldots, \mathbf{x}_{N_u}^T\right]^T, \tag{5.11}$$

where, $\mathbf{x}_i$, $i = 1, \ldots, N_u$, is the signal vector for the i-the user. Also, $\mathbf{w} = \left[\mathbf{w}_1^T, \ldots, \mathbf{w}_{N_u}^T\right]^T \sim \mathcal{CN}(\mathbf{0}, \sigma_{\mathbf{w}}^2\mathbf{I})$ is a $N_u N_r \times 1$ vector that represents the white Gaussian noise. In more detail, $\mathbf{w}_i$, $i = 1, \ldots, N_u$, is the Gaussian noise observed by the i-th user. In this way, by taking into

account (5.1), (5.3), (5.7), and (5.10), the received signal at the i-th user is given as:

$$\mathbf{y}_i = \breve{\mathbf{H}}_i \breve{\mathbf{P}}_i \breve{\mathbf{D}}_i \mathbf{x}_i + \breve{\mathbf{H}}_i \sum_{\substack{k=0 \\ k \neq i}}^{N_u} \breve{\mathbf{P}}_k \breve{\mathbf{D}}_k \mathbf{x}_k + \mathbf{w}_i, \;\; i = 1, \ldots, N_u. \tag{5.12}$$

In the right side of (5.12), the first term represents the intended received signal to the i-th user, while the second term denotes the interference caused by the transmission to the other users. The precoding method of interest in this section is ZF. The main characteristic of ZF precoding is the total elimination of interference between different users and between different antennas of the same user. Hence, ZF precoding is an efficient method that can be used for the formation of a MU architecture based on MSR-SM. Therefore, if the channel matrix is expressed as in (5.4), the ZF precoding matrix is expressed as:

$$\breve{\mathbf{P}} = \mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H \right)^{-1} \mathbf{\Xi}^{-\frac{1}{2}}. \tag{5.13}$$

Note that the deployment of ZF precoding requires that: $N_t \geq N_u N_r$.

Let the ZF precoder in (5.13), the diagonal normalization matrix in (5.10) can be re-written as:

$$\breve{\mathbf{D}} = \mathbf{\Xi}^{\frac{1}{2}} \mathbf{D}_{\text{MU}}, \tag{5.14}$$

where, $\mathbf{D}_{\text{MU}} = \mathbf{diag}\left(d_1, \ldots, d_{N_u N_r}\right)$ is a diagonal matrix. The i-th element of the main diagonal of $\mathbf{D}_{\text{MU}}$ is equal to:

$$d_i = \sqrt{\frac{1}{\left[\left(\mathbf{H}\mathbf{H}^H \right)^{-1} \right]_{i,i}}}, \; i = 1, \ldots, N_u N_r. \tag{5.15}$$

In order to gain a better understanding of the received signal of each user, (5.4), (5.13), and (5.14) can be plugged into (5.1). By doing so, the collective received signal is expressed as:

$$\mathbf{y} = \mathbf{\Xi}^{\frac{1}{2}} \mathbf{D}_{\text{MU}} \mathbf{x} + \mathbf{w}. \tag{5.16}$$

Alternatively, (5.16) can be written as:

$$\mathbf{y}_i = \sqrt{\xi_i} \mathbf{D}_i \mathbf{x}_i + \mathbf{w}_i, \;\; i = 1, \ldots, N_u, \tag{5.17}$$

since $\mathbf{D}_{\text{MU}} = \mathbf{diag}\left(\mathbf{D}_1 \ldots, \mathbf{D}_{N_u}\right)$ is a block diagonal matrix, where $\mathbf{D}_i$ is defined as the i-th $N_r \times N_r$ block matrix of the main block diagonal of $\mathbf{D}_{\text{MU}}$.

From (5.17), it is highlighted that the choice of the transmitted vector $\mathbf{x}_i$, $i = 1, \ldots, N_u$, determines the way that information is transmitted to each user. For example, conventional SMX transmission is obtained if all the elements of $\mathbf{x}_i$ are drawn from a conventional M-ary constellation diagram $\mathcal{M}$. On the other hand, MSR-SM can be obtained by appropriately choosing the symbol vectors $\mathbf{x}_i$, $i = 1, \ldots, N_u$, as discussed in Chapter 4.2.2.

At the users' side, the reconstruction of the transmitted bit-streams is undertaken by detecting the transmitted vectors $\mathbf{x}_i$, $i = 1, \ldots, N_u$. Provided that the i-th user is aware of ξ_i and $\mathbf{D}_i$, this can be implemented at every user independently by using the following ML detector:

$$(\tilde{\mathbf{x}}_i) = \arg\min_{\mathbf{x}_i} \|\mathbf{y}_i - \sqrt{\xi_i}\mathbf{D}_i\mathbf{x}_i\|_2^2, \;\; i = 1, \ldots, N_u. \tag{5.18}$$

5.3 Computational Complexity Analysis

The analysis of the computation complexity of the new architecture is presented in this section. Also, the computational complexity of the new architecture with respect to the corresponding complexity of the conventional MIMO broadcast channel based on ZF precoding and SMX are discussed.

Considering (4.6), it is revealed that the sparsity of the transmission alphabet $\mathcal{B}_i$ of MSR-SM can be used in order to offer a low computational complexity at the transmitter. Provided that the precoding matrix $\breve{\mathbf{P}}$ of (5.13) is precomputed offline before transmission, the transmitted signal $\mathbf{s} = \breve{\mathbf{P}}\breve{\mathbf{D}}\mathbf{x}$ in (5.1) can be computed with $\mathcal{C}_t = N_t(8N_uN_s - 2) + 2N_uN_s$ real operations (additions or multiplications). It is clear that as N_s takes lower values, the complexity of the transmitter $\mathcal{C}_t$ is also reduced. In the extreme case, where N_s is equal to N_r, which is its maximum possible value, SMX transmission is obtained. In this case, the computational complexity of $\mathcal{C}_t$ is maximized. Therefore, it can be concluded that MU MSR-SM offers lower computational complexity at the transmitter than the conventional spatially multiplexed MIMO broadcast channel. For both schemes, in fact, the precoding matrix $\breve{\mathbf{P}}$ is precomputed with the same computational complexity. In addition, both schemes require N_t RF front-ends at the transmitter, since they both rely on linear precoding.

At the users' side, the computational complexity of the ML detector of MSR-SM, as given in (5.18), is higher compared with the corresponding complexity of the conventional MIMO broadcast channel. This occurs because the conventional spatially multiplexed MIMO broadcast channel decouples the N_r data streams of each user. In this way, a per stream ML detector can be employed at each receive antenna of each user. Hence, the detection complexity of the conventional MIMO broadcast channel is lower compared to the complexity of (5.18). However, the authors of [21] propose a suboptimal detector for MSR-SM which achieves almost the same computational complexity as the previous decoupled detector. More details are given in [21]. The study of the detector of [21] is, however, outside of the scope of this section.

5.4 Evaluation of the Average Bit Error Probability

In this section, the ABEP of each individual user and of the whole MU system are derived. The ABEP of the i-th user, $P_{\text{bit}}^{i}(\gamma)$, for a given transmit SNR γ, can be bounded as follows:

$$P_{\text{bit}}^{i}(\gamma) \leq \frac{1}{|\mathcal{B}_i| k_{\text{MSR-SM}}} \sum_{\mathbf{x}_i} \sum_{\substack{\hat{\mathbf{x}}_i \\ \hat{\mathbf{x}}_i \neq \mathbf{x}_i}} d(\mathbf{x}_i \rightarrow \hat{\mathbf{x}}_i) P_{\text{e}}^{i}(\mathbf{x}_i \rightarrow \hat{\mathbf{x}}_i, \gamma), \tag{5.19}$$

using the union bound technique. In (5.19), $P_{\text{e}}^{i}(\mathbf{x}_i \rightarrow \hat{\mathbf{x}}_i, \gamma)$ represents the PEP of transmitting $\mathbf{x}_i$ to the i-th user while the detector decides in favor of the erroneous symbol vector $\hat{\mathbf{x}}_i$. The number of different bits between the bit-words represented by $\mathbf{x}_i$ and $\hat{\mathbf{x}}_i$ is denoted by $d(\mathbf{x}_i \rightarrow \hat{\mathbf{x}}_i)$. Furthermore, $|\mathcal{B}_i| = M^{N_s} 2^{\left\lfloor \log_2 \binom{N_r}{N_s} \right\rfloor}$ denotes the number of all possible transmitted symbol vectors to the i-th user.

The evaluation of (5.19) requires the knowledge of $P_{\text{e}}^{i}(\mathbf{x}_i \rightarrow \hat{\mathbf{x}}_i, \gamma)$, which is the expectation of the instantaneous PEP over all channel realizations. Let the detector of the i-th user in (5.18), a symbol error occurs at this user when:

$$\mathcal{E}_i\left(\mathbf{x}_i, \hat{\mathbf{x}}_i\right) = \left\{ \|\mathbf{y}_i - \sqrt{\xi_i}\mathbf{D}_i\mathbf{x}_i\|_2^2 > \|\mathbf{y}_i - \sqrt{\xi_i}\mathbf{D}_i\hat{\mathbf{x}}_i\|_2^2 \right\}. \tag{5.20}$$

The previous equation can be further simplified as:

$$\mathcal{E}_i\left(\mathbf{x}_i, \hat{\mathbf{x}}_i\right) = \left\{ -\text{Re}\{\mathbf{c}_i^H \mathbf{D}_i \mathbf{w}_i\} > \frac{\sqrt{\xi_i}\mathbf{c}_i^H \mathbf{D}_i^2 \mathbf{c}_i}{2} \right\}, \tag{5.21}$$

using some algebraic manipulations. In (5.21), the vector $\mathbf{c}_i$ is defined as $\mathbf{c}_i = \mathbf{x}_i - \hat{\mathbf{x}}_i$. Given

the statistical distribution of $\mathbf{w}_i$, it can be shown that:

$$-\text{Re}\{\mathbf{c}_i^H \mathbf{D}_i \mathbf{w}_i\} \sim \mathcal{N}\left(0, \frac{\sigma_{\mathbf{w}_i}^2}{2}\sqrt{\xi_i}\mathbf{c}_i^H \mathbf{D}_i^2 \mathbf{c}_i\right). \tag{5.22}$$

Hence, the instantaneous PEP of the i-th user (conditioned on $\mathbf{D}_i$) is given as:

$$\begin{aligned} P_{\text{e}}^i\left(\mathbf{x}_i \to \hat{\mathbf{x}}_i, \gamma | \mathbf{D}_i^2\right) &= P\left(-\text{Re}\{\mathbf{c}_i^H \mathbf{D}_i \mathbf{w}_i\} > \frac{\xi_i \mathbf{c}_i^H \mathbf{D}_i^2 \mathbf{c}_i}{2}\right) \\ &= Q\left(\sqrt{\frac{\mathbf{c}_i^H \mathbf{D}_i^2 \mathbf{c}_i}{2}\xi_i \gamma}\right), \end{aligned} \tag{5.23}$$

where, in the last line, the identity $\frac{1}{\sigma_{\mathbf{w}_i}^2} = \gamma$ is used, which follows from the definition of γ.

For notational convenience, the following variables are defined:

$$z_i = \mathbf{c}_i^H \mathbf{D}_i^2 \mathbf{c}_i, \tag{5.24}$$

and

$$\breve{\gamma}_i = \xi_i \gamma. \tag{5.25}$$

In addition, using the tight upper bound of the Q-function given in (4.24), (5.23) can be simplified (bounded) as:

$$P_{\text{e}}^i\left(\mathbf{x}_i \to \hat{\mathbf{x}}_i, \breve{\gamma}_i | z_i\right) \leq \frac{1}{6} e^{-z_i \breve{\gamma}_i} + \frac{1}{12} e^{-\frac{z_i}{2}\breve{\gamma}_i} + \frac{1}{4} e^{-\frac{z_i}{4}\breve{\gamma}_i}. \tag{5.26}$$

From (5.26), it can be observed that the average over all possible realization of the RV z_i has to be evaluated. This is expressed as:

$$\begin{aligned} P_{\text{e}}^i\left(\mathbf{x}_i \to \hat{\mathbf{x}}_i, \breve{\gamma}_i\right) \leq \frac{1}{6}\text{E}_{z_i}\left[e^{-z_i \breve{\gamma}_i}\right] + \frac{1}{12}\text{E}_{z_i}\left[e^{-\frac{z_i}{2}\breve{\gamma}_i}\right] \\ + \frac{1}{4}\text{E}_{z_i}\left[e^{-\frac{z_i}{4}\breve{\gamma}_i}\right]. \end{aligned} \tag{5.27}$$

Thus, the evaluation of the expectation in (5.27) requires the PDF of the RV z_i. To this end,

using an algebraic elaboration on (5.24), the RV z_i can be re-written as:

$$z_i = \sum_{k=1}^{N_r} |x_k - \hat{x}_k|^2 d_k^2 = \sum_{x_k - \hat{x}_k \neq 0} |x_k - \hat{x}_k|^2 d_k^2, \tag{5.28}$$

where, x_k and $\hat{x}_k$, $k = 1, \ldots, N_r$, are the k-th elements of $\mathbf{x}_i$ and $\hat{\mathbf{x}}_i$, $i = 1, \ldots, N_u$, respectively. Furthermore, d_k, $k = 1, \ldots, N_r$, is the k-th element of the main diagonal of $\mathbf{D}_i$.

Usually, in the literature, the RVs d_k^2 are assumed to be statistically independent in order to simplify the mathematical analysis [180, 181]. This assumption is, however, in contradiction with the structure of $d_k^2 = 1/\left[\left(\mathbf{H}_i\mathbf{H}_i^H\right)^{-1}\right]_{k,k}$. In fact, the realization of every RV d_k^2 occurs using the same mathematical operations on the same random matrix $\mathbf{H}_i$. More specifically, the following holds:

$$d_k^2 = \frac{1}{\left[\text{adj}\left(\mathbf{H}_i\mathbf{H}_i^H\right)\right]_{k,k}} \det\left(\mathbf{H}_i\mathbf{H}_i^H\right), \tag{5.29}$$

where, adj$(\cdot)$ is the adjoint matrix and $\det(\cdot)$ is the matrix determinant. The observation of (5.29) shows that for different values of $k = 1, \ldots, N_r$, the realization of the RVs d_k^2 affects one another, since they are produced via the same mathematical formula using the same random elements of $\mathbf{H}_i$. This implies that the RVs d_k^2 are dependent. An empirical confirmation for the previous argument can be obtained by computing the Pearson product-moment correlation coefficient between any pair of the previous RVs using multiple samples. In this way, it can be shown that the Pearson product-moment correlation coefficient takes non-zero values. The analytical evaluation of these correlation coefficients is difficult to be obtained since it requires the joint PDF between each pair of the RVs d_k^2. An additional confirmation is provided in Section 5.6.1, where the empirical PDF of z_i is depicted against the theoretical PDF derived below. Hence, it can be concluded that the RVs d_k^2 are statistically dependent.

Due to the dependence between the RVs d_k^2, a different approach, compared to the state-of-the-art literature, is proposed: it is taken into account that d_k^2, $k = 1, \ldots, N_r$, are dependent and correlated gamma RVs.

For notational convenience, the variables

$$b_j = |c_k|^2 = |x_k - \hat{x}_k|^2, \tag{5.30}$$

and

$$X_j = d_k^2, \tag{5.31}$$

$j = 1, \ldots, N_i$, are introduced only for those values of k in (5.28) for which it holds that $c_k = x_k - \hat{x}_k \neq 0$. Here, c_k is the k-th element of $\mathbf{c}_i$. In addition, N_i is the number of non-zero elements of $\mathbf{c}_i$. Thus, the value of N_i depends on the considered pair of $\mathbf{x}_i$ and $\hat{\mathbf{x}}_i$. Therefore, (5.28) can be re-written as:

$$z_i = \sum_{j=1}^{N_i} b_j X_j = \sum_{j=1}^{N_i} Z_j. \tag{5.32}$$

In (5.32), Z_j is defined as:

$$Z_j = b_j X_j. \tag{5.33}$$

The PDF of d_k^2 is explicitly derived in [42] as a gamma distribution with:

$$X_j = d_k^2 \sim \text{Gamma}\left(L_{\text{MU}}, 1\right) \tag{5.34}$$

and

$$L_{\text{MU}} = N_t - N_u N_r + 1. \tag{5.35}$$

Therefore, the RVs $Z_j = b_j X_j$, $j = 1, \ldots, N_j$, are distributed as:

$$Z_j \sim \text{Gamma}\left(L_{\text{MU}}, b_j\right), \tag{5.36}$$

with a PDF given by:

$$f_{Z_j}(x) = \frac{1}{b_j^{L_{\text{MU}}} \Gamma(L_{\text{MU}})} x^{L_{\text{MU}}-1} e^{\frac{x}{b_j}} H_0(x). \tag{5.37}$$

Since X_j, $j = 1, \ldots, N_j$, are correlated RVs, also, $Z_j = b_j X_j$, $j = 1, \ldots, N_j$ are correlated RVs. This implies that z_i is a RV which is equal to the sum of correlated Gamma RVs. For this reason and based on [182, Corollary 1], the PDF of z_i is given as:

$$f_{z_i}(x) = \left[\prod_{l=1}^{N_i} \left(\frac{\breve{\alpha}_1}{\breve{\alpha}_l}\right)^{L_{\text{MU}}}\right] \left[\sum_{k=0}^{+\infty} \frac{\breve{\delta}_k x^{N_i L_{\text{MU}}+k-1} e^{-\frac{x}{\breve{\alpha}_1}}}{\breve{\alpha}_1^{N_i L_{\text{MU}}+k} \Gamma\left(N_i L_{\text{MU}} + k\right)}\right] H_0(x). \tag{5.38}$$

In (5.38), $\breve{\alpha}_l$, $l = 1, \ldots, N_i$, are the eigenvalues of:

$$\breve{\mathbf{A}} = \breve{\mathbf{B}}\breve{\mathbf{R}} \tag{5.39}$$

in ascending order, and $\breve{\mathbf{B}}$ is the diagonal matrix

$$\breve{\mathbf{B}} = \mathbf{diag}\left(b_1, \ldots, b_{N_i}\right), \tag{5.40}$$

where b_l, $l = 1, \ldots, N_i$, is the square of the absolute value of the l-th non-zero element of $\mathbf{c}_i$. Also, $\breve{\mathbf{R}}$ is a $N_i \times N_i$ matrix defined as:

$$\breve{\mathbf{R}} = \begin{bmatrix} 1 & \sqrt{\rho_c} & \cdots & \sqrt{\rho_c} \\ \sqrt{\rho_c} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \sqrt{\rho_c} \\ \sqrt{\rho_c} & \cdots & \sqrt{\rho_c} & 1 \end{bmatrix}, \tag{5.41}$$

where, ρ_c is the Pearson product-moment correlation coefficient between any pair of two different RVs of the main diagonal of $\mathbf{D}_i^2$. The structure of d_k^2 shows that the Pearson product-moment correlation coefficient between every pair of two different RVs of the main diagonal of $\mathbf{D}_i^2$ takes the value of ρ_c. Finally, $\breve{\delta}_k$, $k = 0, 1, 2, \ldots$, are given as:

$$\breve{\delta}_{k+1} = \begin{cases} 1, & k = -1, \\ \frac{k}{k+1} \sum_{i=1}^{k+1} \left[\sum_{j=1}^{N} \left(1 - \frac{\breve{\alpha}_1}{\breve{\alpha}_j}\right)^i \right] \breve{\delta}_{k+1-i}, & k = 0, 1, 2, \ldots. \end{cases} \tag{5.42}$$

Let the PDF of z_i given in (5.38), the evaluation of (5.27) over all possible realizations of z_i can be performed by evaluating expectations of the following form:

$$g\left(y\right) = \mathrm{E}_{z_i}\left[e^{-y\breve{\gamma}_i z_i}\right]. \tag{5.43}$$

The evaluation of the previous expectation is given in Appendix C. Since $\breve{\gamma}_i = \xi_i \gamma$ and using

the result from Appendix C, the PEP of the i-th user is given as:

$$P_{\mathrm{e}}^{i}(\mathbf{x}\rightarrow\hat{\mathbf{x}},\gamma)\leq\frac{\left[\prod_{l=1}^{N_i}\left(\frac{\breve{\alpha}_1}{\breve{\alpha}_l}\right)^{L_{\mathrm{MU}}}\right]}{6}\left(\breve{\alpha}_1\xi_i\gamma+1\right)^{-N_iL_{\mathrm{MU}}}\sum_{k=0}^{+\infty}\breve{\delta}_k\left(\breve{\alpha}_1\xi_i\gamma+1\right)^{-k}$$
$$+\frac{\left[\prod_{l=1}^{N_i}\left(\frac{\breve{\alpha}_1}{\breve{\alpha}_l}\right)^{L_{\mathrm{MU}}}\right]}{12}\left(\frac{\breve{\alpha}_1}{2}\xi_i\gamma+1\right)^{-N_iL_{\mathrm{MU}}}\sum_{k=0}^{+\infty}\breve{\delta}_k\left(\frac{\breve{\alpha}_1}{2}\xi_i\gamma+1\right)^{-k}$$
$$+\frac{\left[\prod_{l=1}^{N_i}\left(\frac{\breve{\alpha}_1}{\breve{\alpha}_l}\right)^{L_{\mathrm{MU}}}\right]}{4}\left(\frac{\breve{\alpha}_1}{4}\xi_i\gamma+1\right)^{-N_iL_{\mathrm{MU}}}\sum_{k=0}^{+\infty}\breve{\delta}_k\left(\frac{\breve{\alpha}_1}{4}\xi_i\gamma+1\right)^{-k}. \quad (5.44)$$

Thus, the evaluation of the ABEP of the i-th user follows from (5.19), by using (5.44).

In addition to the performance of each user, the whole system performance is of interest. A metric for evaluating the whole system performance is the system ABEP. Assuming that the detection process at each user is performed independently, the system ABEP is expressed as:

$$P_{\mathrm{bit}}^{\mathrm{System}}(\gamma)=\frac{1}{N_u}\sum_{i}^{Nu}P_{\mathrm{bit}}^{i}(\gamma_i). \quad (5.45)$$

An upper bound of (5.45) can be obtained by using the upper bound of the PEP, $P_{\mathrm{bit}}^{i}(\gamma_i)$, of each user given in (5.19).

5.5 Analysis of Diversity Order and Coding Gain

In the high SNR regime, the user and system performance can be characterized in terms of diversity order and coding gain [173]. Initially, these performance measures are given for the i-th user, $i=1,\ldots,N_u$ and then the analysis is generalized from the system standpoint based on (5.45).

In order to compute the diversity order and coding gain of the i-th user, a high SNR approximation for the PEPs of the i-th user is needed. By using mathematical steps similar to Section 5.4 and based on the Chernoff bound of the Q-function:

$$Q\left(x\right)\leq\frac{1}{2}e^{-\frac{x^2}{2}}, \quad (5.46)$$

the PEP of the i-th user can be bounded as:

$$P_{\mathrm{e}}^{i}(\mathbf{x}_i \to \hat{\mathbf{x}}_i, \gamma) \leq \frac{\left[\prod_{l=1}^{N_i} \left(\frac{\breve{\alpha}_1}{\alpha_l}\right)^{L_{\mathrm{MU}}}\right]}{2} \left(\frac{\breve{\alpha}_1}{4}\xi_i\gamma + 1\right)^{-N_i L_{\mathrm{MU}}} \times \sum_{k=0}^{+\infty} \breve{\delta}_k \left(\frac{\breve{\alpha}_1}{4}\xi_i\gamma + 1\right)^{-k}. \tag{5.47}$$

From (5.47), as proved in Appendix D, a high SNR ($\gamma \to +\infty$) approximation of the PEP of the i-th user can be formulated as:

$$P_i^{+\infty}(\mathbf{x}_i \to \hat{\mathbf{x}}_i, \gamma) \lessapprox \left[\frac{\breve{\alpha}_1 \sqrt[N_i L_{\mathrm{MU}}]{2}}{4}\xi_i\gamma\right]^{-N_i L_{\mathrm{MU}}} + \mathrm{o}\left(\gamma^{-N_i L_{\mathrm{MU}}}\right). \tag{5.48}$$

Then, a high SNR approximation of the SER of the i-th user is obtained as follows:

$$\mathrm{SER}_i^{+\infty} \lessapprox \frac{1}{|\mathcal{B}|} \sum_{\mathbf{x}_i} \sum_{\substack{\hat{\mathbf{x}}_i \\ \hat{\mathbf{x}}_i \neq \mathbf{x}_i}} P_i^{+\infty}(\mathbf{x}_i \to \hat{\mathbf{x}}_i, \gamma), \tag{5.49}$$

In (5.49), $\mathcal{B}$ denotes the set of all possible transmitted symbol vectors to a generic user.

It can be observed that the high SNR approximation of the SER in (5.49) is a linear combination of $P_i^{+\infty}(\mathbf{x}_i \to \hat{\mathbf{x}}_i, \gamma)$, as given in (5.48), for all possible pairs of $\mathbf{x}_i$ and $\hat{\mathbf{x}}_i$. Therefore, as $\gamma \to +\infty$, the slope of (5.49) is determined by the smallest exponent of γ in (5.48), i.e. $N_i L_{\mathrm{MU}}$. The smallest value of $N_i L_{\mathrm{MU}}$ occurs when $N_i = 1$. In fact, the dominant addends of (5.49) are those for which $N_i = 1$. Therefore, the high SNR approximation of the SER in (5.49) can be further approximated by considering only these dominant addends. In addition, the careful observation of (5.48) shows that the matrix $\breve{\mathbf{A}}$ in (5.39) reduces to a scalar if $N_i = 1$. This implies $\breve{\alpha}_1 = b_1$, where b_1 is given in (5.30).

With this simplification, a more insightful approximation of (5.49) can be obtained. More specifically, from (5.48), (5.49) can be expressed as:

$$\mathrm{SER}_i^{+\infty} \lessapprox \frac{1}{|\mathcal{B}|} \sum_{\mathbf{x}_i} \sum_{n=1}^{N_s} \sum_{x \in \mathcal{M}} \sum_{\substack{\hat{x} \in \mathcal{M} \\ \hat{x} \neq x}} \left[\frac{|x - \hat{x}|^2 \sqrt[L_{\mathrm{MU}}]{2}}{4}\xi_i\gamma\right]^{-L_{\mathrm{MU}}} + \sum_{k=2}^{N_s} \mathrm{o}\left(\gamma^{-kL_{\mathrm{MU}}}\right). \tag{5.50}$$

Using a similar approach as [183, Chapter 5.2.9], an upper bound for (5.50) is obtained by retaining only the minimum distance, denoted by $d_{\min}$, between every pair of the constellation points $\{x, \hat{x}\} \in \mathcal{M}$. By doing so, after some algebraic manipulations of the summations in (5.50), the following high SNR approximation of the SER of the i-th user is obtained:

$$\mathrm{SER}_i^{+\infty} \lessapprox \left[\frac{d_{\min}^2}{4} \sqrt[L_{\mathrm{MU}}]{\frac{2}{N_s \left(M^2 - M\right)}} \xi_i \gamma \right]^{-L_{\mathrm{MU}}} + \sum_{k=2}^{N_s} \mathrm{o}\left(\gamma^{-k L_{\mathrm{MU}}}\right). \tag{5.51}$$

The bound in (5.51) may be loose for high values of the constellation order M [183, Chapter 5.2.9]. However, it is conveniently formulated for providing insightful information on the achievable diversity order and coding gain. If M is large, if needed, a tighter bound may be obtained by following the guidelines in [183, Chapter 5.2.9].

Based on the definitions of the diversity order and coding gain available in [173], (5.51) reveals that the diversity order of the i-th user is:

$$d_i = L_{\mathrm{MU}}, \tag{5.52}$$

and that the corresponding coding gain is:

$$c_i = \frac{d_{\min}^2}{4} \sqrt[L_{\mathrm{MU}}]{\frac{2}{N_s \left(M^2 - M\right)}} \xi_i. \tag{5.53}$$

Since $L_{\mathrm{MU}} = N_t - N_u N_r + 1$, from (5.52) it follows that the diversity order of the i-user does not depend on the large scale channel effect, but only on the system size (the number of transmit antennas N_r, the number of users N_u, and the number of receive antennas per user N_r). In contrast, (5.53) shows that the coding gain of the i-th user depends on the system size, the number of parallel data streams N_s, the constellation size $\mathcal{M}$ (via $d_{\min}$ and M), and ξ_i which represents the large scale channel effect. Here, it is indirectly assumed that the large scale effect is deterministic.

From the system-level standpoint, the diversity order can be computed by approximating the

system SER for high SNR ($\gamma \to +\infty$) as follows:

$$\begin{aligned}\mathrm{SER}_{\mathrm{System}}^{+\infty} &= \frac{1}{N_u}\sum_{i=1}^{N_u}\mathrm{SER}_i^{+\infty} \\ &\lessapprox \frac{1}{N_u}\sum_{i=1}^{N_u}\left[\frac{d_{\min}^2}{4}\sqrt[L_{\mathrm{MU}}]{\frac{2}{N_s\left(M^2-M\right)}}\xi_i\gamma\right]^{-L_{\mathrm{MU}}} \\ &\quad + \sum_{k=2}^{N_s}\mathrm{o}\left(\gamma^{-kL_{\mathrm{MU}}}\right), \end{aligned} \tag{5.54}$$

where, the last step in (5.54) exploits the high SNR approximation in (5.51). In order to express (5.54) in a convenient form that explicitly provides information on the diversity order and coding gain, an upper bound based on the smallest value of ξ_i, $i = 1, \ldots, N_u$ is used. More specifically, the following holds:

$$\begin{aligned}\mathrm{SER}_{\mathrm{System}}^{+\infty} &\lessapprox \left[\frac{d_{\min}^2}{4}\sqrt[L_{\mathrm{MU}}]{\frac{2}{N_s\left(M^2-M\right)}}\xi_{\min}\gamma\right]^{-L_{\mathrm{MU}}} \\ &\quad + \sum_{k=2}^{N_s}\mathrm{o}\left(\gamma^{-kL_{\mathrm{MU}}}\right), \end{aligned} \tag{5.55}$$

where, $\xi_{\min} = \min\left(\xi_1, \ldots, \xi_{N_u}\right)$. From (5.55) and [173], it follows that the system diversity order is:

$$d_{\mathrm{System}} = L_{\mathrm{MU}}, \tag{5.56}$$

and the corresponding coding gain is:

$$c_{\mathrm{System}} = \frac{d_{\min}^2}{4}\sqrt[L_{\mathrm{MU}}]{\frac{2}{N_s\left(M^2-M\right)}}\xi_{\min}. \tag{5.57}$$

Comparing (5.56) and (5.57) with (5.52) and (5.53), respectively, it is concluded that the diversity order and the coding gain from the user and system standpoints are the same. The main difference is that the system-level coding gain in (5.57) is dominated by the large scale channel effect of the user having the weakest channel, i.e., the smallest value of ξ_i, $i = 1, \ldots, N_u$.

Based on the obtained expressions of the diversity order and coding gain, the proposed transmission scheme can be compared against the conventional MIMO broadcast channel. Note that the proposed mathematical framework is directly applicable to the conventional MIMO broadcast channel by simply setting $N_s = N_r$. Therefore, (5.52), (5.53), (5.56), and (5.57) can be

directly used for comparing MU MSR-SM and the conventional MIMO broadcast channel.

From (5.52) and (5.56), in particular, it is concluded that the diversity order is independent of N_s. As a result, both schemes achieve the same diversity order. As for the conventional MIMO broadcast channel, this conclusion is in agreement with the results available in [70]. This further validates the correctness of the new mathematical framework.

The comparison of (5.53) and (5.57), on the other hand, shows that the coding gain depends on N_s. Therefore, the coding gain of MU MSR-SM, where, in general, $N_s < N_r$ holds, is different from the coding gain of the conventional MIMO broadcast channel, where $N_s = N_r$. Since both schemes offer the same diversity order, the scheme providing the highest coding gain also results in the lowest BER. Hence, the superiority of a scheme compared to the other can be assessed by a direct inspection of the following coding gain ratio:

$$\begin{aligned}\psi_i &= \frac{c_i^{\mathrm{MSR-SM}}}{c_i^{\mathrm{SMX}}} \\ &= \left(\frac{d_{\min}^{\mathrm{MSR-SM}}}{d_{\min}^{\mathrm{SMX}}}\right)^2 \sqrt[L_{\mathrm{MU}}]{\frac{N_r\left(M_{\mathrm{SMX}}^2 - M_{\mathrm{SMX}}\right)}{N_s\left(M_{\mathrm{MSR-SM}}^2 - M_{\mathrm{MSR-SM}}\right)}}.\end{aligned} \tag{5.58}$$

By appropriately choosing the constellation orders $M_{\mathrm{MSR-SM}}$ and M_{SMX} for MU MSR-SM and for the conventional MIMO broadcast channel, respectively, the same spectral efficiency can be guaranteed. In (5.58), the coding gain of the i-th user of MU MSR-SM is denoted by $c_i^{\mathrm{MSR-SM}}$ and the coding gain of the same user in the conventional MIMO broadcast channel is denoted by c_i^{SMX}. Furthermore, $d_{\min}^{\mathrm{MSR-SM}}$ and $d_{\min}^{\mathrm{SMX}}$ denote the minimum distance between every pair of points of the adopted signal constellations for MU MSR-SM and for the conventional MIMO broadcast channel, respectively. From (5.58), it follows that MU MSR-SM performs better than the conventional MIMO broadcast channel if $\psi_i > 1$. If $M_{\mathrm{MSR-SM}} = M_{\mathrm{SMX}}$, a direct inspection of (5.58) reveals that $\psi_i > 1$ and that it increases as N_s decreases. As a result, in this case, MU MSR-SM outperforms the conventional MIMO broadcast channel. In general, however, it holds that $M_{\mathrm{MSR-SM}} \neq M_{\mathrm{SMX}}$. In this case, a direct analysis of (5.58) is more difficult. The ratio ψ_i can, however, be numerically computed. Table 6.1 provides typical values of ψ_i in dB scale, by assuming the same spectral efficiency for both schemes. Table 6.1 shows that, for the considered system setups, MU MSR-SM provides a higher coding gain than the conventional MIMO broadcast channel if $N_s = N_r - 1$. More specifically, the coding gain is in the range 1 – 2.49 dB. If $N_s < N_r - 1$, on the other hand, the conventional MIMO

System Configuration	N_s	k_{user} (bpsp)	ψ_i (in dB)
$N_t = 20$, $N_r = 4$, $N_u = 4$	3	8	1.49
$N_t = 20$, $N_r = 4$, $N_u = 4$	2	8	-2.49
$N_t = 20$, $N_r = 4$, $N_u = 4$	1	8	-11.05
$N_t = 20$, $N_r = 5$, $N_u = 4$	4	10	1.93
$N_t = 16$, $N_r = 4$, $N_u = 4$	3	8	2.49

Table 5.1: *Coding gain of MU MSR-SM, based on (5.58), with respect to the conventional MIMO broadcast channel ($N_s = N_r$).*

broadcast channel provides a higher coding gain. As a result, (5.58) can be used for the system optimization and for ensuring that MSR-SM is superior to the SotA.

5.6 Simulation Results and Discussion

The objective of this section is twofold. Firstly, to validate the theoretical results of Sections 5.4 and 5.5 using simulation results. Secondly, to provide a performance comparison between MU MSR-SM and the benchmark MIMO broadcast channel. In this latter case, in particular, $N_s = N_r$ is assumed and no SM is used. In all of the studied scenarios, the number of users is equal to $N_u = 4$. As described in Section 5.2, the wireless channel of the i-th user, $i = 1, \ldots, N_u$, is generated following a complex Gaussian distribution $\left(\breve{\mathbf{H}}_i \sim \mathcal{CN}(\mathbf{0}, \xi_i \mathbf{I})\right)$. More specifically, ξ_i, $i = 1, \ldots, N_u$, is set equal to 1, 0.75, 0.5, and 0.25, for user 1, 2, 3, and 4, respectively. These values demonstrate the large scale channel effect of different users. Values of ξ_i which are close to one model strong channels, while as ξ_i is reduced and approaches zero, less strong channels are modeled. For a fair comparison, the M-ary constellations of both schemes are normalized such that $\mathrm{E}_{\mathbf{x}}[\mathbf{x}] = 1$ and $\mathrm{E}_{\mathbf{x}_i}[\mathbf{x}_i] = \frac{1}{N_u}$. Hence, the transmit SNR of the whole system is $\gamma = \frac{1}{\sigma^2_{\mathbf{w}}}$.

5.6.1 Validation of (5.38)

In Section 5.4, the ABEP of MU MSR-SM is derived by using the PDF of z_i given in (5.38). The derivation of (5.38) is based on the fact that the RVs d_k^2, $k = 1, \ldots, N_r$, are statistically correlated. In order to confirm this, Fig. 5.2 illustrates the empirical PDF of (5.32) and compares it against its analytical expression in (5.38). In addition, Fig. 5.2 shows the analytical PDF of (5.32) under the incorrect assumption that d_k^2, $k = 1, \ldots, N_r$, are statistically independent RVs, as usually considered in the literature for mathematical tractability. If this assumption was valid, the PDF of (5.32) could be directly obtained by using the result from [182, Theorem 1].

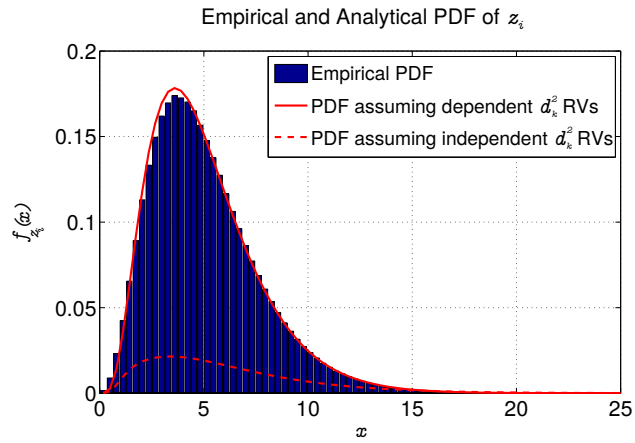


Figure 5.2: *Illustration of empirical and analytical PDF of (5.32) by assuming that: i) the RVs* d_k, $k = 1, \ldots, 2$ *are statistically dependent and ii) they are independent. Setup:* $\mathbf{H} \sim \mathcal{CN}(\mathbf{0}_{2\times 4}, \mathbf{I}_{2\times 4})$*; and ii)* $b_1 = 0.5$ *and* $b_1 = 1.2$.

Note that the integration, form zero to infinity, of both theoretical PDFs is equal to 1 [184]. From Fig. 5.2, it is observed that the theoretical PDF of (5.32) perfectly matches its empirical PDF. In contrast, when the RVs d_k^2, $k = 1, \ldots, N_r$ are assumed to be independent, the obtained PDF from [182, Theorem 1] deviates from the empirical results.

5.6.2 Validation of the Theoretical Analysis

The upper bounds derived in Section 5.4 are compared against Monte Carlo simulations in Figs. 5.3 and 5.4. Furthermore, the same figures illustrate the upper bounds of the ABEP for the system and for each user when the high SNR approximation of the PEP in (5.48) is used. As shown in Figs. 5.3 and 5.4, the analytical bounds of the ABEP are tight in the high SNR, both for (5.44) and (5.48). More specifically, in the high SNR, the analytical results can be considered as an excellent approximation of the simulation results. In contrast, in the low SNR, there is a small difference between the theoretical and simulation results. However, this is a well known phenomenon that originates from using union bound methods [183]. Finally, the diversity order and coding gain analysis of Section 5.5 is also verified from Figs. 5.3 and 5.4. In more detail, the slope of the simulated BER curves of each user is $L_{\text{MU}} = N_t - N_u N_r + 1$. Note that, in high SNR, both the simulated BER curves and the theoretical curves are parallel to the asymptotic of $\text{SNR}^{-L_{\text{MU}}}$ (in the logarithmic scale). This confirms the validity of the derived diversity order. In addition, the simulated curves show that the behavior of the coding gain of the i-th user depends on ξ_i. Similarly, the simulated curves show that the system coding gain is

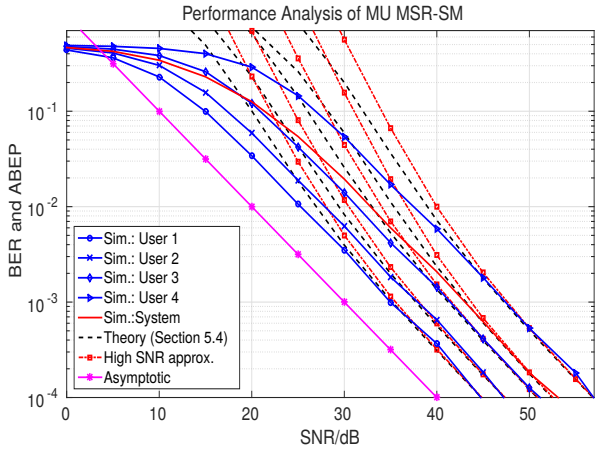


Figure 5.3: *Performance analysis of MU R-SM ($N_s = 1$) for four users, when ZF with P-CSIT is employed: simulation results vs. the bounds in Section 5.4. The high SNR approximation of the ABEP is calculated using the PEP given in (5.48). Setup: $N_t = 16$, $N_r = 4$, ξ_i, $i = 1, \ldots, N_u$, takes the values 1, 0.75, 0.5, and 0.25 for the user 1, 2, 3, and 4, respectively.*

dominated by $\xi_{\min}$. These conclusions are in perfect agreement with the analysis presented in Section 5.5.

5.6.3 Analysis of the Impact of N_s

Fig. 5.5 shows the BER of MU MSR-SM as a function of N_s. It shows, in particular, that the BER of each user gets better as N_s increases. This happens because higher values of N_s require a lower modulation order of QAM in order to achieve the same spectral efficiency. Table I, however, shows that the optimal value of N_s is not necessarily equal to N_r, i.e., the conventional MIMO broadcast channel.

5.6.4 BER Comparison with the Conventional Broadcast Channel

Fig. 5.6 shows the BER of MU MSR-SM and the conventional MIMO broadcast channel. In the low SNR, conventional MIMO offers a slightly better BER than the new scheme. In the high SNR, on the other hand, the new architecture outperforms the benchmark. In particular, MU MSR-SM provides 1 dB gain at BER=10^{-4}. This finding is in agreement with the mathematical analysis of the ratio of the coding gains in (5.58).

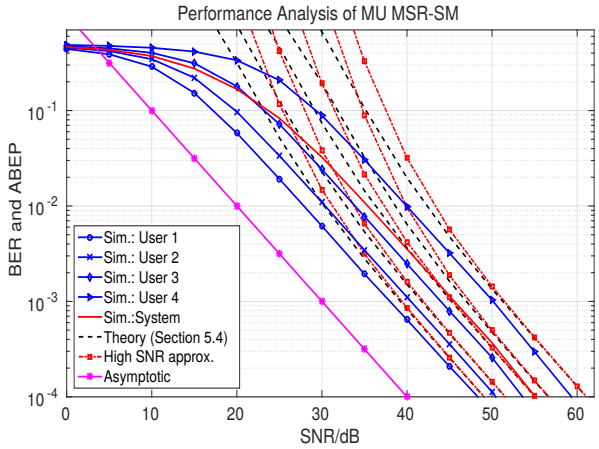


Figure 5.4: *Performance analysis of MU MSR-SM ($N_s = 2$) for four users, when ZF with P-CSIT is employed: simulation results vs. the bounds in Section 5.4. The high SNR approximation of the ABEP is calculated using the PEP given in (5.48). Setup: $N_t = 16$, $N_r = 4$, ξ_i, $i = 1, \ldots, N_u$, takes the values 1, 0.75, 0.5, and 0.25 for the user 1, 2, 3, and 4, respectively.*

Configuration	k_{system} (bpsp)	Target BER	RAER [%]
$N_t = 20, N_r = 4, N_u = 4$	32	10^{-1}	-20%
$N_t = 20, N_r = 4, N_u = 4$	32	10^{-2}	1%
$N_t = 20, N_r = 4, N_u = 4$	32	10^{-4}	18%

Table 5.2: *RAER of MU MSR-SM with respect to the conventional MIMO broadcast channel.*

5.6.5 Energy Efficiency Comparison with the Conventional Broadcast Channel

Table 5.2 presents the system RAER of MU MSR-SM with respect to the conventional MIMO broadcast channel, by assuming the same system setup as in Section 5.6.4. Note that the definition of the metric of RAER is given in (4.92). Both schemes provide the same BER and the same spectral efficiency. If BER=10^{-1}, the conventional MIMO broadcast channel is more energy efficient. For practical values of the BER less than 10^{-2}, on the other hand, MU MSR-SM becomes more energy efficient. For example, an energy efficiency gain of 18% is achieved by MU MSR-SM at BER=10^{-4}.

5.7 Summary

In this chapter, the incorporation of MSR-SM for application to the MIMO broadcast channel was introduced and its BER performance was mathematically studied. More specifically, based on the union bound technique, an accurate mathematical framework for its performance eval-

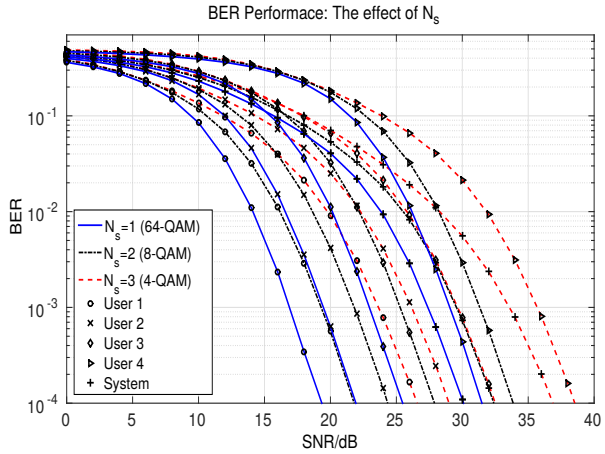


Figure 5.5: *BER performance of MU MSR-SM as a function of N_s. Setup: $N_t = 20$, $N_r = 4$, $N_u = 4$. The spectral efficiency is 8 bpcu.*

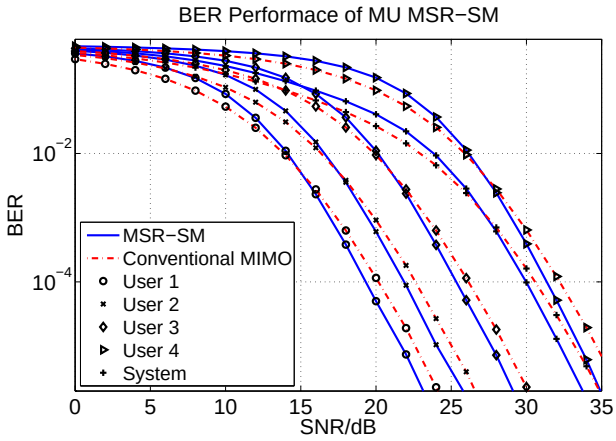


Figure 5.6: *BER performance of MU MSR-SM versus the conventional MIMO broadcast channel. Setup: $N_t = 20$, $N_r = 4$, $N_u = 4$. The spectral efficiency is 8 bpcu by using 4-QAM for both schemes.*

uation was proposed and discussed. From this framework, it is proved that MSR-SM provides the same diversity order as the conventional MIMO broadcast channel, while offering a better coding gain in the high SNR regime. Also, this performance gain is achieved with a reduction of the complexity of the transmitter. Numerical simulations were shown in order to substantiate the gain predicted by the analysis. As a result, MSR-SM is shown to be a promising transmission scheme for the MIMO broadcast channel.

Chapter 6
Dual-Hop Hybrid-Spatial Modulation: The Extension of Spatial Modulation in a Distributed Framework

6.1 Introduction

The aim of this chapter is to propose and study an architecture which is suitable for the downlink. Motivated by the concept of VMIMO and the potential of SM/R-SM, Chapter 6 proposes a non-cooperative relay architecture based on SM/R-SM. Using a half-duplex DF protocol, a scheme that achieves information conveyance using R-SM [137] in the first hop and SM [15] in the second hop is proposed. In this way, an architecture called DH-HSM is formed. The term Hybrid Spatial Modulation (HSM) stems from the fact that the new architecture employs R-SM in the first hop and SM in the second hop. The RNs are able to employ a Centralized Detection (CD) or a Distributed Detection (DD) algorithm, depending on their ability to coordinate. In this way, DH-HSM conveys information by extending the novel transmission mechanism of SM in a distributed framework.

In order to practically solve or reduce the problem of the activation of multiple RNs, without affecting the complexity of the RNs or the DN, the transmission mechanism of the first hop is carefully designed. More specifically, under the presence of CSIT, MIMO linear precoding is utilized. In this way, the effect of the wireless channel is practically eliminated or reduced. Furthermore, significant receive SNR gains are obtained at the RNs.

The evaluation of DH-HSM against the SotA is conducted using the metric of the BER. The employed benchmark systems are the: i) single relay; ii) best relay selection; and iii) distributed STBC dual hop architectures; and iv) the direct communication link. Even though the comparison of DH-HSM with the best relay selection and distributed STBC is not absolutely fair for DH-HSM, the new architecture is able to achieve significant BER gains in the majority of the system setups. It is shown that the gain of DH-HSM over the best relay selection scheme can reach up to 9 dB for BER=10^{-4} and a very large scale system setup. In order to validate the

obtained simulation results, the ABEP of the new architecture is analyzed using the well known union bound method [169]. Note that the derivation of the ABEP is conducted both for the CD and DD algorithms at the RNs. Finally, the performance of the new architecture is studied under the presence of realistic I-CSIT at the SN.

6.2 System Model

Fig. 6.1 shows the system model of the proposed DH-HSM architecture. It considers a BS equipped with N_t antennas which acts as the SN. Furthermore, it is assumed that the new architecture includes N_R single antenna RNs connected with an error-free backhaul link. This assumption is realistic in scenarios where a backhaul link can be established, such as attaching the RNs to a building, ship, plane, or train. In this case, one of the RN acts as the main processing unit. Note that this assumption also corresponds to the scenario where there is one RN with multiple antennas. However, this assumption is not realistic for a general cellular system where the multiple RNs and the DN correspond to mobile terminals. Thus, the system model described in this section acts either as a reference scenario or as one of the application scenarios previously described. *Later, in Section 6.3, the new architecture is extended in a distributed framework where there is no backhaul connection between the RNs*. The DN can be considered as a single or multi-antenna node (with N_D antennas).

In addition, it is assumed that there is no direct link between the SN and the DN due to the poor channel condition. The SN acquires CSIT using either the reciprocity principle when it is applicable or using a low-rate feedback link from the relays. The DN acquires CSI using a training sequence transmitted from the relays. Especially, when TDD is employed, DH-HSM is able to supply the SN and the DN with CSI by transmitting a single training sequence from the RNs. For the sake of system presentation, in Section 6.2, it is assumed that the SN possesses P-CSIT.

The new architecture is configured with $N_t \geq N_R$ and N_R equal to a power of 2. Under the assumptions that the transmitter employs a linear precoder and the wireless channel is flat quasi-static, the system equation of the first hop is expressed in a matrix form as:

$$\mathbf{y}_{\mathrm{R}} = \mathbf{H}_{\mathrm{SR}}\mathbf{P}\mathbf{D}\mathbf{x} + \mathbf{w}_{\mathrm{SR}}. \tag{6.1}$$

In (6.1), $\mathbf{y}_{\mathrm{R}}$ represents the $N_R \times 1$ received signal at the RNs in a vector form. Also, $\mathbf{H}_{\mathrm{SR}} \in$

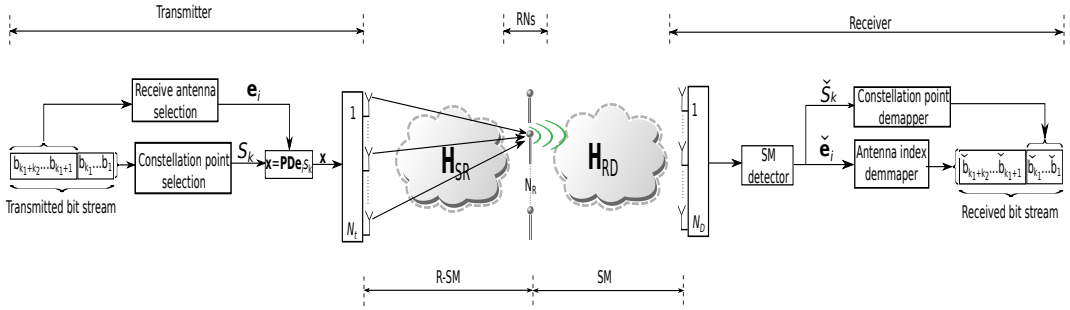


Figure 6.1: *System Model of DH-HSM.*

$\mathcal{C}^{N_R \times N_t}$ represents the channel matrix which is assumed to distribute as $\mathbf{H}_{\text{SR}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$. Similar to Chapter 4, $\mathbf{P}$ denotes the $N_t \times N_R$ linear precoding matrix and $\mathbf{D}$ is the normalization matrix which is used to ensure that the transmitted power is not amplified by the precoder. The designs of $\mathbf{P}$ and $\mathbf{D}$ are presented in Section 4.2.1. Furthermore, $\mathbf{x}$ represents the transmitting symbol vector. In this chapter, the elements of $\mathbf{x}$ are set such that $\mathrm{E}_{\mathbf{x}}\left[\|\mathbf{x}\|_2^2\right] = 1$. Finally, $\mathbf{w}_{SR} \in \mathcal{C}^{N_D}$ denotes the i.i.d. additive complex Gaussian noise with $\mathbf{w}_{SRi} \sim \mathcal{CN}(0, \sigma^2_{w_{SR}})$.

In order to establish R-SM transmission from the SN to the RNs, such as [20, 137], $\mathbf{x}$ is defined as in (4.4) in Chapter 4.2.1. In this way, $k_1 + k_2$ bits are transmitted from the SN to the RNs in a symbol period (first hop).

Under the assumption of ZF precoding using P-CSIT at the SN, the observation at each of the RNs is given as:

$$\mathbf{y}_R = \mathbf{Dx} + \mathbf{w}_{SR}. \tag{6.2}$$

If the sparsity of $\mathbf{x}$ is taken into account, (6.2) can be written as:

$$\begin{aligned} y_{Rj} &= d_j s_k + w_{SRj}, \quad j = i, \\ y_{Rj} &= w_{SRj}, \qquad\quad j \neq i, \end{aligned} \tag{6.3}$$

where $j = 1, ..., N_R$.

Due to the fact that the RNs communicate via a backhaul-link, the following centralized low-complexity ML detector from [137] can be employed:

$$(\hat{i}, \hat{s}_k) = \arg\min_{i, s_k} d_i^2 |s_k|^2 - 2 d_i \mathrm{Re}\{y_{Ri}^* s_k\}. \tag{6.4}$$

As it can be seen from (6.4), the CSI that the R-SM detector requires is d_i, $i = 1, \ldots, N_R$.

Note that (6.4) ensures the activation of a single RN during the second hop.

During the second slot, a SM-like transmission mechanism is applied where only the receiving RN from the previous slot is activated. This RN acts as a single antenna in a conventional co-located SM system. In this way, the system equation of the RNs-DN link is given as:

$$\mathbf{y}_D = \mathbf{H}_{RD}\mathbf{x}(\hat{i}, \hat{s}_k) + \mathbf{w}_{RD}, \tag{6.5}$$

where $\mathbf{H}_{\text{RD}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ stands for the $N_D \times N_R$ RNs-DN wireless channel. In addition, $\mathbf{x}(\hat{i}, \hat{s}_k) = \mathbf{e}_{\hat{i}}\hat{s}_k$ is the $N_R \times 1$ vectorized detected symbol from the previous slot. Also, $\mathbf{w}_{RD} \in \mathcal{C}^{N_D}$ is the i.i.d. additive complex Gaussian noise with $\mathbf{w}_{RDi} \sim \mathcal{CN}(0, \sigma^2_{w_{RD}})$. Finally, the received signal at the DN is denoted by $\mathbf{y}_D$. The optimal ML detection for SM is given in [104] as:

$$(\check{i}, \check{s}_k) = \arg\min_{i, s_k} \|\mathbf{y}_D - \mathbf{H}_{RD}\mathbf{e}_i s_k\|_2^2. \tag{6.6}$$

Note that the detector in (6.6) can be further simplified if the sparsity of $\mathbf{e}_i$ is taken into account. In [104], the ML detector for SM is expressed as:

$$(\check{i}, \check{s}_k) = \arg\min_{i, s_k} \|\mathbf{h}^i_{RD}c\|_2^2 - 2\text{Re}\{\mathbf{y}_D \left(\mathbf{h}^i_{RD}\right)^H s_k\}, \tag{6.7}$$

where, $\mathbf{h}^i_{RD}$ is the i-th, $i = 1, \ldots, N_R$, column of $\mathbf{H}_{RD}$. The computational analysis of the detection process of (6.7) is provided in [104]. In this paper, it is shown that the computational complexity of (6.7) is approximately $2N_rN_R + N_RM + M$ real multiplications, where M is the order of the deployed constellation.

Hence, in this way, a dual-hop architecture which conveys $k_t = \frac{k_1+k_2}{2}$ bits per signaling period (bpsp) is formed. Note that k_1+k_2 is divided by two, because the transmitted bitstream requires two symbol periods to reach the DN.

A special case of the above transmission mechanism, called Dual Dual Hop-Hybrid Space Shift Keying (DH-HSSK), is formed if it holds that $s_k = 1$. In such a scenario, the information is conveyed using only the index i of the receiving RN.

6.3 Extension to a Distributed and Uncoordinated Architecture

In some scenarios there is no backhaul-link between the RNs. This is the case when: i) the delay requirement does not allow communication between the RNs; ii) the cost of installing a backhaul connection is high or iii) the RNs are remote terminals (like cell phones). In this section, the proposed centralized relay architecture from Section 6.2 is extended to a distributed one with the employment of a decentralized detection algorithm at the RNs. All the other configurations of the proposed architecture remain the same.

During the first hop, the received signal at the j-th RN is described in (6.3). It is straightforward to see that if it holds that $s_0 = 0$, (6.3) can be easily interpreted as the following $(M+1)$-ary Hypothesis Test (HT):

$$H_i: \quad y_{Rj} = d_j s_i + w_{SRj}, \quad i = 0, ..., M, \tag{6.8}$$

which can be independently employed at each RN. In (6.8), s_i takes values from $\{s_0, s_1, ..., s_M\}$. Clearly, the H_0 hypothesis ($s_0 = 0$) corresponds to the case where the j-th RN is not the receiving node in the first hop. Thus, this RN remains silent during the second hop. On the other hand, all the other $H_1, ..., H_M$ hypotheses correspond to the case where the j-th node is the receiving/activating RN and the relayed conventional symbol is $s_1, ..., s_M$, respectively.

In this case, the distributed ML detector at each RN is reduced to the usual minimum distance rule:

$$(\hat{s}_i) = \arg \min_{s_i \in \{s_0, s_1, ..., s_M\}} |y_{Rj} - d_j s_k|. \tag{6.9}$$

Given that $d_j s_k$ is precomputed and stored in the memory before the beginning of transmission, the computational complexity of (6.9) at each RN is just one complex subtractions per iteration.

Clearly, the independent execution of the detector of (6.9) at each RN could result into the activation of multiple RNs during the second hop. In the other extreme case, the distributed detectors of (6.9) of all RNs could detect the zero symbol s_0. When this extreme case occurs, all the RNs remain silent and no symbol is relayed to the DN. Due to the fact that the DN is unaware of this situation, its detector decides randomly based only on the ever present Gaussian noise.

In the literature, these problems are treated using error-aware detection at the DN [151, 154], which increases the complexity of the DN. In order to keep the complexity of the DN low,

the phenomenon of the activation of multiple RNs is decided to be ignored. However, as it is demonstrated in Section 6.5, the new DD algorithm benefits from the linear precoding at the SN, which results in high SNR gains at the RNs. Thus, the probability of activating multiple RNs is reduced. Because of this, practically and in most scenarios, either there is no performance difference from the CD algorithm, or there is a small BER penalty.

6.4 Theoretical Evaluation of the Average Bit Error Probability

In this section, the union bound technique [169] is deployed in order to provide bounds on the ABEP of the proposed architecture. This is undertaken for both the CD (Section 6.2) and DD (Section 6.3) algorithms in the RNs. In order to provide tractability to the theoretical analysis, ZF precoding with P-CSIT at the SN is considered. This decision is taken because it is well known that the ABEP performance analysis of MMSE or any other regularized ZF-like precoding method is a challenging problem [70]. Besides, in high SNRs, ZF precoding approaches the performance of MMSE precoding [72, 185].

6.4.1 Average Bit Error Probability When The Relay Nodes Employ Centralized Detection

When the CD algorithm of (6.4) is employed at the RNs, DH-HSSK/HSM is a regenerative relay system with coordinated detection. Thus, a similar procedure as [150, 186] can be followed for the derivation of the ABEP. In this way, the end-to-end ABEP is expressed as:

$$P_{\mathrm{CD}}(\gamma_1, \gamma_2) = P_1(\gamma_1) + P_2(\gamma_2) - P_1(\gamma_1)P_2(\gamma_2). \tag{6.10}$$

In (6.10), γ_1 and γ_2 denotes the SNRs of the first and the second hop, respectively. Also, $P_{\mathrm{CD}}(\gamma_1, \gamma_2)$ is the end-to-end ABEP, given γ_1 and γ_2. Furthermore, $P_1(\gamma_1)$ and $P_2(\gamma_2)$ represents the ABEP of the first and second hop, respectively. It is not difficult to see that $P_1(\gamma_1)$ and $P_2(\gamma_2)$ are the ABEP of R-SM and SM, respectively. The ABEP of SM is widely studied in the literature [14, 104, 187]. However, the ABEP of R-SM is a challenging problem, due to the use of ZF precoding [70].

The derivation of the ABEP of R-SM is conducted in detail in Section 4.3.1. More specifically, the ABEP of R-SM is given from (4.8). Furthermore, the PEP for each type of symbol error

(signal, space and joint error) is given from: i) (4.28) for a signal error; ii) (4.38) for a space error; and iii)(4.44) for a joint error. Note that, just like SM, in R-SM there are three possible types of symbol errors. A signal error:

$$\mathcal{E}_1 = \{\mathbf{x} \to \hat{\mathbf{x}} | \{s_k \to \hat{s}_k, \mathbf{e}_i \to \mathbf{e}_i\}\} \tag{6.11}$$

occurs when the receiving antenna (RN) (represented by $\mathbf{e}_i$) is correctly detected, while the transmitted symbol s_k is incorrectly detected . In contrast, a space error:

$$\mathcal{E}_2 = \left\{\mathbf{x} \to \hat{\mathbf{x}} | \left\{s_k \to s_k, \mathbf{e}_i \to \mathbf{e}_{\hat{i}}\right\}\right\} \tag{6.12}$$

occurs when s_k is correctly detected and the receiving antenna (RN) is incorrectly detected. Finally, the incorrect detection of s_k and $\mathbf{e}_i$ creates a joint symbol error:

$$\mathcal{E}_3 = \left\{\mathbf{x} \to \hat{\mathbf{x}} | \left\{s_k \to \hat{s}_k, \mathbf{e}_i \to \mathbf{e}_{\hat{i}}\right\}\right\}. \tag{6.13}$$

In this chapter, and in alignment with [179], instead of using the bounds of (4.28), (4.38), and (4.44), simpler bounds for each type of symbol errors are provided without proof. In fact, the provided bounds can be obtained following a similar procedure as Section 4.3.1. The derivation of the new bounds begins with the assessment of ABEP of the first hop (ABEP of R-SM). Therefore, the PEP of transmitting $\mathbf{x}$ at the RN and receiving $\hat{\mathbf{x}}$ at the RNs has to be evaluated. This is undertaken for all possible combinations of

$$\mathbf{x}, \hat{\mathbf{x}} \in \mathcal{A}_1 = \{\mathbf{e}_i s_k | i \in \{1, \ldots, N_R\} \cap k \in \{1, \ldots, M\}\} \tag{6.14}$$

and $\mathbf{x} \neq \hat{\mathbf{x}}$. Given that (6.2) holds and the ML detector selects the symbol $\mathbf{x}$ which minimizes the quantity $\|\mathbf{y}_R - \mathbf{D}\mathbf{x}\|_2^2$, the instantaneous PEP conditioned on $\mathbf{D}^2$ is expressed as:

$$P_{\mathrm{R-SM}}(\mathbf{x} \to \hat{\mathbf{x}} | \mathbf{D}^2) = Q\left(\sqrt{\frac{\mathbf{c}^H \mathbf{D}^2 \mathbf{c}}{2}\gamma_1}\right), \tag{6.15}$$

where $\mathbf{c} = \mathbf{x} - \hat{\mathbf{x}}$ and $\gamma_1 = 1/\sigma^2_{\mathbf{w}_{SR}}$. The focus here is on deriving the PEP by averaging (6.15) over all possible realizations of $\mathbf{D}^2$. If the same procedure as Section 4.3.1 is followed; but instead using the bound of (4.24) for the $Q(\cdot)$ function, the Chernoff bound $Q(x) \leq \frac{1}{2}e^{-\frac{x^2}{2}}$ [169] is deployed; simpler bounds for the PEPs for each type of symbol error (signal, space,

or joint error) can be obtained. The new bounds can be found below. Note that the proof of these bounds is omitted since it follows the same reasoning as Section 4.3.1. Thus, the PEP of a signal error $\mathcal{E}_1$ is shown to be bounded as:

$$P_{\mathrm{R-SM}}(\mathbf{x} \to \hat{\mathbf{x}}|\mathcal{E}_1) \leq \frac{1}{2}\left[\frac{|\delta|^2}{4}\gamma_1 + 1\right]^{-L}. \tag{6.16}$$

Note that (6.16) is also the PEP of a MISO system with ZF precoding. As it can be seen, this MISO system has a diversity order of $L = N_t$ (given that there is one receive antenna). This conclusion agrees with the indirect diversity analysis of [70]. Furthermore, following the same procedure as Section 4.3.1, it can be shown that PEP of a space error $\mathcal{E}_2$ has the following bound:

$$P_{\mathrm{R-SM}}(\mathbf{x} \to \hat{\mathbf{x}}|\mathcal{E}_2) \leq \frac{(1-\rho_c)^{-L}}{2t^{2L}}\ {}_1F_0\left(L;\ ;\frac{\rho_c}{(1-\rho_c)^2t^2}\right), \tag{6.17}$$

where,

$$t = \frac{|s_k|^2\gamma_1}{4} + \frac{1}{1-\rho_c}. \tag{6.18}$$

Similarly, the upper bound of a joint error is given as:

$$P_{\mathrm{R-SM}}(\mathbf{x} \to \hat{\mathbf{x}}|\mathcal{E}_3) \leq \frac{[(1-\rho_c)t_1t_2]^{-L}}{2}\ {}_1F_0\left(L;\ ;\frac{\rho_c}{t_1t_2(1-\rho_c)^2}\right),$$

where,

$$t_1 = \frac{|s_k|^2\gamma_1}{4} + \frac{1}{1-\rho_c} \tag{6.19}$$

and

$$t_2 = \frac{|\hat{s}_k|^2\gamma_1}{4} + \frac{1}{1-\rho_c}. \tag{6.20}$$

All the other parameters in (6.16), (6.17), and (6.19) are defined just like in (4.28), (4.38), and (4.44), respectively.

In the final step of the proof, the bounds of (6.16), (6.17), and (6.19) are deployed in order to express the ABEP of the first hop as:

$$P_1(\gamma_1) \leq \frac{1}{Mk_t}\sum_{\mathbf{x}}\sum_{\substack{\hat{\mathbf{x}}\\ \hat{\mathbf{x}}\neq\mathbf{x}}} d(\mathbf{x} \to \hat{\mathbf{x}})P_{\mathrm{R-SM}}(\mathbf{x} \to \hat{\mathbf{x}}), \tag{6.21}$$

where, it is reminded that $d(\mathbf{x} \to \hat{\mathbf{x}})$ denotes the number of different bits between the bit sequences represented by $\mathbf{x}$ and $\hat{\mathbf{x}}$.

In order to complete the derivation of (6.10), the ABEP of the second hop is expressed using the results from [187]. Hence, the ABEP of the second hop is bounded as:

$$P_2(\gamma_2) \leq \frac{1}{Mk_t} \sum_{\mathbf{x}} \sum_{\substack{\hat{\mathbf{x}} \\ \hat{\mathbf{x}} \neq \mathbf{x}}} d(\mathbf{x} \to \hat{\mathbf{x}}), P_{\mathrm{SM}}(\mathbf{x} \to \hat{\mathbf{x}}), \tag{6.22}$$

where,

$$P_{\mathrm{SM}}(\mathbf{x} \to \hat{\mathbf{x}}) = \frac{1}{2} \det \left(\mathbf{I}_{\zeta,\zeta} + \gamma_2 \frac{\Lambda}{\sqrt{2}} \right). \tag{6.23}$$

Here, it holds that that $\zeta = N_D N_R$ and $\Lambda = \mathbf{I}_{N_D,N_D} \otimes \mathbf{c}\mathbf{c}^H$.

6.4.2 Average Bit Error Probability When The Relay Nodes Employ Distributed Detection

In this section, the ABEP of DH-HSSK/HSM is provided when the RNs utilize the DD of (6.9). In this type of uncoordinated detection, the methodology that is followed in Section 6.4.1 does not hold, because during the second hop multiple RNs can be active. In this case, the transmitted symbol vector of the second hop is not a legal SM symbol. Thus, (6.10) is not valid.

In order to overcome this problem, the union bound technique is used in order to express the end-to-end ABEP of DH-HSM as:

$$P_{\mathrm{DD}}(\gamma_1, \gamma_2) \leq \frac{1}{Mk_t} \sum_{\mathbf{x}} \sum_{\substack{\hat{\mathbf{x}} \\ \hat{\mathbf{x}} \neq \mathbf{x}}} d(\mathbf{x} \to \hat{\mathbf{x}}) P_{\mathrm{DD}}(\mathbf{x} \to \hat{\mathbf{x}}, \gamma_1, \gamma_2), \tag{6.24}$$

where $P_{\mathrm{DD}}(\mathbf{x} \to \hat{\mathbf{x}}, \gamma_1, \gamma_2)$ is the end-to-end PEP of transmitting $\mathbf{x}$ at the SN and detecting erroneously $\hat{\mathbf{x}} \neq \mathbf{x}$ at the DN, given γ_1 and γ_2 (the SNR of each hop). Note that $\mathbf{x}, \hat{\mathbf{x}} \in \mathcal{A}_1 = \{\mathbf{e}_i s_k | i \in \{1, \ldots, N_R\} \cap k \in \{1, \ldots, M\}\}$.

It is not difficult to see that, using the total probability theorem, the PEP of an end-to-end symbol error is expressed as:

$$P_{\mathrm{DD}}(\mathbf{x} \to \hat{\mathbf{x}}, \gamma_1, \gamma_2) = \sum_{\tilde{\mathbf{x}} \in \mathcal{A}} P_{\mathrm{SR}}\left(\mathbf{x} \to \tilde{\mathbf{x}}, \gamma_1\right) P_{\mathrm{RD}}\left(\tilde{\mathbf{x}} \to \hat{\mathbf{x}}, \gamma_2\right), \tag{6.25}$$

where $\tilde{\mathbf{x}} \in \mathcal{A}$ are all the possible detected/transmitted symbol vectors at the RNs. The probability that $\mathbf{x}$ is transmitted at the SN and $\tilde{\mathbf{x}}$ is detected/retransmitted at the RNs is denoted as

$P_{\mathrm{SR}}\left(\mathbf{x} \to \tilde{\mathbf{x}}, \gamma_1\right)$. Furthermore, $P_{\mathrm{RD}}\left(\tilde{\mathbf{x}} \to \hat{\mathbf{x}}, \gamma_2\right)$ represents the probability that $\tilde{\mathbf{x}}$ is transmitted at the RNs and $\hat{\mathbf{x}}$ is detected at the DN.

Before continuing, the set $\mathcal{A} = \{\mathcal{A}_0, \mathcal{A}_1, \ldots, \mathcal{A}_{N_R}\}$ is defined. Here, it holds that $\mathcal{A}_0 = \{\mathbf{0}_{N_R,1}\}$. The subset $\mathcal{A}_0$ corresponds to the very special case where all the RNs remain silent during the second hop (for more details, the reader is referred to Section 6.3). Every other subset $\mathcal{A}_q$, with $q = 1, \ldots, N_R$ and $|\mathcal{A}_q| = \binom{N_R}{q} M^q$, contains all the possible vectors, of size $N_R \times 1$, with exactly q non-zero elements. Their non-zero elements take values from $\mathcal{M} = \{s_1, \ldots, s_k\}$. Mathematically, this is represented as $\mathcal{A}_q = \{\mathbf{e}_{i_1} s_{k_1} + \ldots + \mathbf{e}_{i_q} s_{k_q} | \mathcal{B}_1 \cap \mathcal{B}_2\}$, where $\mathcal{B}_1 = \{\{i_1, \ldots, i_q\} \in \{1, \ldots, N_R\} | \{i_1 \neq \ldots \neq i_q\}\}$, $\mathcal{B}_2 = \{\{k_1, \ldots, k_q\} \in \{1, \ldots, M\}\}$, and $\mathbf{e}_{i_l}$ is the i_l-th column of $\mathbf{I}_{N_R,N_R}$. Note that, clearly, the correct symbol $\mathbf{x}$ belongs to $\mathcal{A}$. Even though the RNs might be able to detect correctly the transmitted vector $\mathbf{x}$, a symbol error could occur during the second hop.

Each probability of the right hand side of (6.25) has to be evaluated. The derivation begins with the evaluation of $P_{\mathrm{SR}}\left(\mathbf{x} \to \tilde{\mathbf{x}}, \gamma_1\right)$. Given that detection is conducted independently at each RN and the Gaussian noise is independent, it holds that:

$$P_{\mathrm{SR}}\left(\mathbf{x} \to \tilde{\mathbf{x}}, \gamma_1\right) = \prod_{j=1}^{N_R} \mathrm{E}_{d_j^2}\left[P_{\mathrm{SR}_j}\left(\mathbf{x}_j \to \tilde{\mathbf{x}}_j, \gamma_1 | d_j^2\right)\right]. \tag{6.26}$$

Here, $\mathbf{x}_j, \tilde{\mathbf{x}}_j \in \{s_0, \mathcal{M}\}$ are the j-th elements of $\mathbf{x}$ and $\tilde{\mathbf{x}}$, respectively. Simple, $\tilde{\mathbf{x}}_j$ belongs to the employed constellation $\mathcal{M}$ or is zero and $\tilde{\mathbf{x}}_j$ is the detected symbol which again belongs to $\mathcal{M}$ or is zero.

Thus, the probability of:

$$P_{\mathrm{SR}_j}\left(\mathbf{x}_j \to \tilde{\mathbf{x}}_j, \gamma_1\right) = \mathrm{E}_{d_j^2}\left[P_{\mathrm{SR}_j}\left(\mathbf{x}_j \to \tilde{\mathbf{x}}_j, \gamma_1 | d_j^2\right)\right] \tag{6.27}$$

has to be evaluated by integrating over all the possible realizations d_j^2. The received signal at the j-th RN is given from (6.3). Hence, given that $\mathbf{x}_j, \tilde{\mathbf{x}}_j \in \{s_0, \mathcal{M}\}$, with $\mathbf{x}_j \neq \tilde{\mathbf{x}}_j$, it can be written that:

$$P_{\mathrm{SR}_j}\left(\mathbf{x}_j \to \tilde{\mathbf{x}}_j, \gamma_1 | d_j^2\right) = Q\left(\sqrt{\frac{\gamma_1 |\mathbf{x}_j - \tilde{\mathbf{x}}_j|^2 d_j^2}{2}}\right). \tag{6.28}$$

For this reason, following the same steps as the derivation of (6.16), it is proven that:

$$P_{\mathrm{SR}_j}\left(\mathbf{x}_j \to \tilde{\mathbf{x}}_j, \gamma_1\right) \leq \frac{1}{2}\left[\frac{|\mathbf{x}_j - \tilde{\mathbf{x}}_j|^2}{4}\gamma_1 + 1\right]^{-L}. \tag{6.29}$$

For the case of $\mathbf{x}_j = \tilde{\mathbf{x}}_j$ (correct transmission to the j-th RN), it holds that:

$$P_{\mathrm{SR}_j}\left(\mathbf{x}_j \to \tilde{\mathbf{x}}_j, \gamma_1\right) = 1 - P_e(\gamma_1) \leq 1 - \sum_{\substack{j=0 \\ \mathbf{x}_j \neq s_j}}^{M} P_{\mathrm{SR}_j}\left(\mathbf{x}_j \to s_j, \gamma_1\right), \tag{6.30}$$

where, $P_e(\gamma_1)$ is the probability of erroneous detection. Hence, the probability of interest is bounded by:

$$P_{\mathrm{SR}_j}\left(\mathbf{x}_j \to \tilde{\mathbf{x}}_j, \gamma_1\right) \leq 1 - \frac{1}{2}\sum_{\substack{j=0 \\ \mathbf{x}_j \neq s_j}}^{M}\left[\frac{|\mathbf{x}_j - s_j|^2}{4}\gamma_1 + 1\right]^{-L}. \tag{6.31}$$

In addition, the evaluation of (6.25) requires the knowledge of $P_{\mathrm{RD}}\left(\tilde{\mathbf{x}} \to \hat{\mathbf{x}}, \gamma_2\right)$. It is known that $\tilde{\mathbf{x}} \in \mathcal{A}$, which means that the RNs retransmit legal or illegal SM symbols. Even in the case of the retransmission of illegal SM symbols, as it is explained in Section 6.3, the DN uses the detector of (6.6). Thus, an error happens when:

$$\mathcal{E}' = \left\{\|\mathbf{y}_D - \mathbf{H}_{RD}\mathbf{x}\|_2^2 > \|\mathbf{y}_D - \mathbf{H}_{RD}\hat{\mathbf{x}}\|_2^2 | \tilde{\mathbf{x}}\right\} \tag{6.32}$$

(given that $\tilde{\mathbf{x}}$ is relayed by the RNs). At this point, $\mathcal{E}'$ can be easily transformed as:

$$\mathcal{E}' = \left\{\|\mathbf{H}_{RD}\mathbf{c}_1 \ + \mathbf{w}_{RD}\|_2^2 > \|\mathbf{H}_{RD}\mathbf{c}_2 \ + \mathbf{w}_{RD}\|_2^2\right\}, \tag{6.33}$$

where, $\mathbf{c}_1 = \tilde{\mathbf{x}} - \mathbf{x}$ and $\mathbf{c}_2 = \tilde{\mathbf{x}} - \hat{\mathbf{x}}$. A further elaboration on $\mathcal{E}'$ results in:

$$\mathcal{E}' = \left\{\|\mathbf{H}_{RD}\mathbf{c}_1\|_2^2 - \|\mathbf{H}_{RD}\mathbf{c}_2\|_2 > 2\mathrm{Re}\{\mathbf{c}^H\mathbf{H}_{RD}^H\mathbf{w}_{RD}\}\right\}, \tag{6.34}$$

where $\mathbf{c} = \mathbf{c}_2 - \mathbf{c}_1$. Given that:

$$-\sum_{k=1}^{N_D}\mathrm{Re}\left\{\mathbf{w}_{RD_k}a_k\right\} \sim \mathcal{N}\left(0, \frac{\sigma_{w_{RD}}^2}{2}\|\mathbf{H}_{RD}\mathbf{c}\|_2^2\right), \tag{6.35}$$

where, $a_k = \sum_{j=1}^{N_D} \mathbf{c}_j^* \left[\mathbf{H}_{RD}\right]_{j,k}^*$, the instantaneous probability of interest can be written as:

$$P_{\mathrm{RD}}\left(\tilde{\mathbf{x}} \to \hat{\mathbf{x}}, \gamma_2 | \mathbf{H}_{RD}\right) = Q\left(\gamma' \frac{\|\mathbf{H}_{RD}\mathbf{c}_2\|_2^2 - \|\mathbf{H}_{RD}\mathbf{c}_1\|_2^2}{\sqrt{\|\mathbf{H}_{RD}\mathbf{c}\|_2^2}}\right), \tag{6.36}$$

where $\gamma' = \sqrt{\gamma_2/2}$. The direct evaluation of the expectation of (6.36), over all channel realizations, is a challenging task. Actually, this expectation requires the PDF of the summation of two Erlang (gamma) RVs divided by a generalized gamma RV, where all RVs are correlated. An approach that gives a tractable solution to the expectation of (6.36) is to use the triangle and inverse triangle inequality and upper and lower bound the RV:

$$Y = \frac{\|\mathbf{H}_{RD}\mathbf{c}_2\|_2^2 - \|\mathbf{H}_{RD}\mathbf{c}_1\|_2^2}{\sqrt{\|\mathbf{H}_{RD}\mathbf{c}\|_2^2}}. \tag{6.37}$$

Unfortunately, this method gives bounds that are loose and does not provide any insight. Therefore, this analysis is not included. An alternative way to evaluate the expectation of (6.36) is numerically. In this way, the proof is completed.

6.5 Results and Discussion

In Section 6.5, theoretical and Monte Carlo simulation results are presented. These results demonstrate the performance of the proposed DH-HSSK/HSM architecture under different forms of CSIT at the SN, both for the CD and DD algorithms.

For the sake of comparison, the new architecture is compared against four benchmark systems: i) the corresponding single relay system; ii) the best relay selection system; iii) a form of distributed Alamouti relaying; and iv) the direct communication link. Note that in all the benchmark systems, the operating principle of the SN and the DN is the same as DH-HSM/ Hybrid Space Shift Keying (HSSK). The SN uses the same linear precoding methods and the DN detects the received signal using ML detection.

The corresponding single relay system is denoted as Dual Hop-with one Relay (DH-w1R). The adoption of this form of comparison for the evaluation of space modulated VMIMO is extensively employed in the literature [26, 148, 150, 151, 154]. The operational principle of DH-w1R is similar to the proposed DH-HSSK/HSM architecture with the following exception. The conveyed constellation point is relayed by a single RN.

It is noted that the RN of DH-w1R detects the relayed constellation point using a ML detector which has almost the same complexity as the ML detector of (6.4) and higher complexity than the detector of (6.9) (all detectors are single stream detectors). Especially, the DD of (6.9) results in a significant lower complexity as the spectral efficiency and the number of the RNs is increased. Because of this, DH-HSSK/HSM can be interpreted as a way of distributing the computational complexity of the single RN of DH-w1R to the multiple RNs of DH-HSSK/HSM. An additional advantage of DH-HSSK/HSM compared with DH-w1R is that, under the assumption of high SNR at the RNs, usually only one out of N_R RNs is re-transmitting the symbol received from the SN. This is very important when the RNs are remote terminals (for example mobile phones) with a battery supply, because a battery life extension is achieved[1]. Hence, if the advantages of DH-HSSK/HSM over DH-w1R are taken into account, with respect to: i) the complexity and power consumption distribution and ii) the BER gains (as shown in the following), it is clear that the cost of having multiple RNs of low complexity and extended battery life compared with the cost of having a single RN of high complexity is insignificant.

The next two benchmark systems that are employed are: i) the best relay selection technique which is denoted as Dual Hop-with Best Relay Selection (DH-wBRS) and ii) the distributed Alamouti code which is denoted as Dual Hop-Distributed Alamouti Code (DH-DAC). In DH-wBRS, a single RN out of N_R is selected to convey the transmitted constellation point in every block of symbols. More specifically, DH-wBRS selects the RN with the most favorable RN-DN channel which has the highest channel gain. The selection of this RN offers the highest instantaneous receive SNR at the DN. In DH-DAC, two symbols are conveyed to two RNs in order to form a distributed Alamouti codeword. The RNs can employ a CD or DD algorithm, just like DH-HSM. In the DD case, the orthogonality of the Alamouti codeword cannot be ensured, whereas in the CD case it can be ensured. Note that DH-DAC requires 4 symbol period in order to convey two constellation points (constructed Alamouti codeword) from the SN to the DN. Clearly, the comparison of DH-HSM with DH-wBRS and DH-DAC is not entirely fair. DH-wBRS uses closed loop transmission during the second hop, while DH-HSM uses open loop transmission. Furthermore, DH-DAC activates two RNs during the second hop, whereas DH-HSM activate only one. Finally, it is stated that: i) DH-DAC is used as a benchmark system only when the RNs are two and ii) no best RN selection criterion is applied in DH-DAC. As shown in the following, even though these comparisons are not entirely

[1]In this work, it is assumed that the energy consumption for the RF power transmission is relatively high compared to the energy consumption of the circuits of the RNs during the detection period.

fair, DH-HSM is able to achieve significant performance gains in the majority of the cases.

The last used benchmark system is the direct SN-DN MIMO communication. In this system, it is assumed that the SN-DN wireless channel is weak and distributes as $\mathcal{CN}(\mathbf{0}, p\mathbf{I})$, where p captures the path loss effect. This assumption is valid because DH-HSM is designed to be applied when the SN-DN channel is weak.

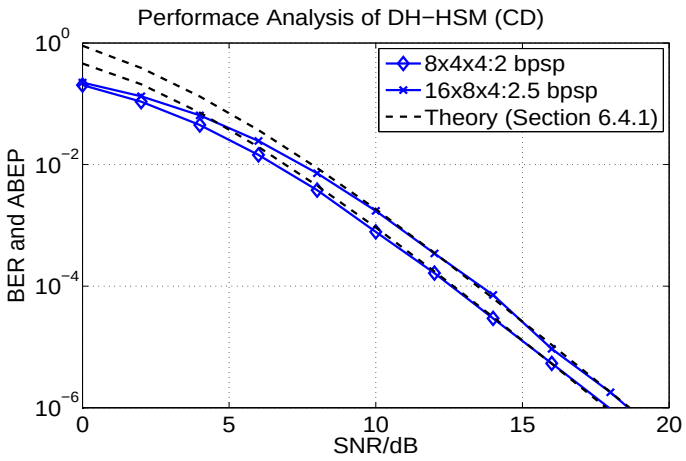
For all systems, the SNR during the first hop is defined as $\gamma_1 = \text{SNR}_{\text{dB}}^{\text{SR}} = 10\log_{10}\frac{P_s^{SR}}{\sigma_{w_{SR}}^2}$, where $P_s^{SR} = \text{E}_{\mathbf{x}}\left[\|\mathbf{PDx}\|_2^2\right]$ is the power transmitted at SN. For a fair comparison between DH-HSSK/HSM and DH-w1R, it is enforced that $P_s = 1$ by using the appropriate: i) normalization matrix $\mathbf{D}$ for the precoder $\mathbf{P}$ (as described in Section 6.2) and ii) the appropriate normalization of the conventional transmitted constellation diagram. During the second hop, the SNR is defined as $\gamma_1 = \text{SNR}_{\text{dB}}^{\text{RD}} = 10\log_{10}\frac{P_s^{RD}}{\sigma_{w_{RD}}^2}$, where $P_s^{RD} = 1$ is the transmitted power from the active RN. Without loss of generality it is assumed that $\gamma_1 = \gamma_2$. Finally, it is assumed that the SN-RNs and RNs-DN wireless channel remain constant over a block symbol. In addition, when I-CSIT is available at the SN, $\bar{\mathbf{H}}_{\text{SR}}$ does not change over a block of symbols.

6.5.1 Confirmation of the Average Bit Error Probability of DH-HSM Using Simulation Results

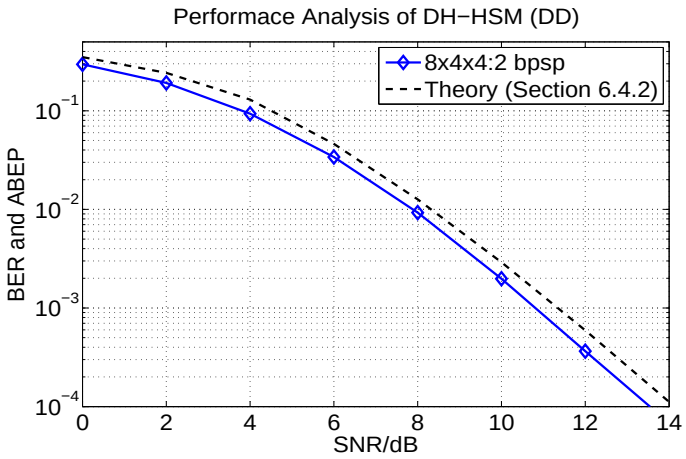
Before proceeding with the comparison of DH-HSM with the benchmark system, Fig. 6.2 shows how close are the obtained simulation results with the theoretical framework of Section 6.4. Fig. 6.2(a) presents the bound of ABEP when CD is employed at the RNs. In addition, Fig. 6.2(b) shows the ABEP bound under DD at the RNs. Two system configurations are considered for CD. The first configuration is a $8 \times 4 \times 4$ ($N_t = 8$, $N_R = 4$, and $N_D = 4$) system with $k_t = 2$ bpsp. Furthermore, the second configuration is a $16 \times 8 \times 4$ system with $k_t = 2.5$ bpsp. As can be seen from Fig. 6.2(a), for CD and in high SNR, the theoretical and simulation curves perfectly match. In low SNR, there is a difference which is a well known phenomenon caused by the union bound technique [169]. In addition, Fig. 6.2(b) demonstrates that the bound obtained for DD in Section 6.4.2 is close to the simulated curve.

6.5.2 Small Scale System Setup

Fig. 6.3 shows the BER performance of DH-HSM when P-CSIT is available at the DN and the system setup is $4 \times 2 \times 4$, with $k_t = 4$ bpsp. For the sake of comparison, Fig. 6.3 includes the



(a) The RNs employ the CD algorithm.



(b) The RNs employ the DD algorithm.

Figure 6.2: *Performance analysis of DH-HSM when ZF with P-CSIT is employed at the RNs, using: i) simulation results and ii) the bounds of Section 6.4.*

performance of: i) DH-w1R; ii) DH-wBRS; iii) DH-DAC; and iii) direct link communication when $p = \{0.1, 1\}$. Especially, the scenario of the direct link communication with a weak SN-DN channel ($p = 0.1$) is a typical application scenario of DH-HSM (or any other dual hop system). Furthermore, as a reference point, the performance of direct link communication when the SN-DN channel is strong with $p = 1$ (same statistics as the SN-RNs and RNs-DN channels) is included.

As it can be seen from Fig. 6.3, DH-HSM offers better BER performance than DH-w1R, DH-wBRS, and DH-DAC due to its multiplexing gain. It is quite notable that in the depicted SNR region, DH-HSM offers better performance than DH-wBRS and DH-DAC, even though, this comparison is not entirely fair. However, as SNR increases to extremely high value,

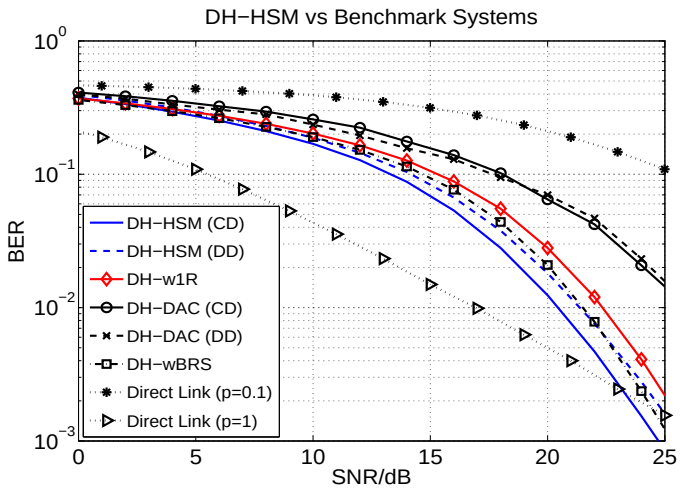


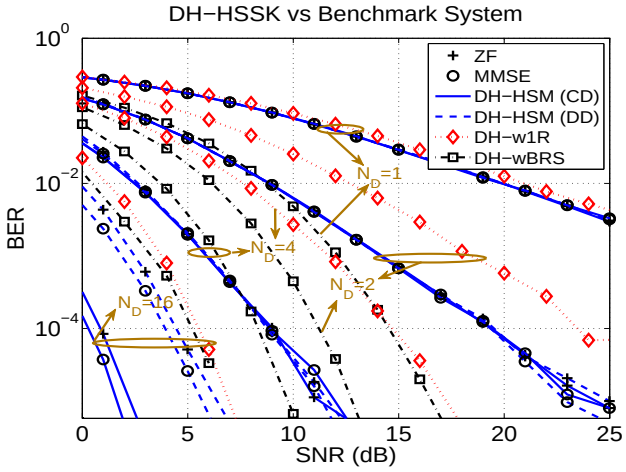
Figure 6.3: *BER performance of a small scale DH-HSM* ($4 \times 2 \times 4$) *system versus DH-w1R, DH-wBRS, DH-DAC, and the* 4×4 *direct link communication when* $p = \{0.1, 1\}$. *The spectral efficiency is* $k_t = 4$ *bpsp.*

DH-wBRS and DH-DAC would offer better performance, due to their diversity achieving techniques employed during the second hop. Note that, just like DH-HSM, the performance of DH-DAC is not affected significantly by the employment of CD or DD at the RNs. Regarding the performance of the direct link communication, it can be seen that when the SN-DN channel is weak ($p = 0.1$), its performance is significantly worse (as expected). For a strong SN-DN channel ($p = 1$), the direct link offers better performance due to its multiplexing gain. However, in high SNR, the direct link is remarkably outperformed due to its unity diversity order [70].

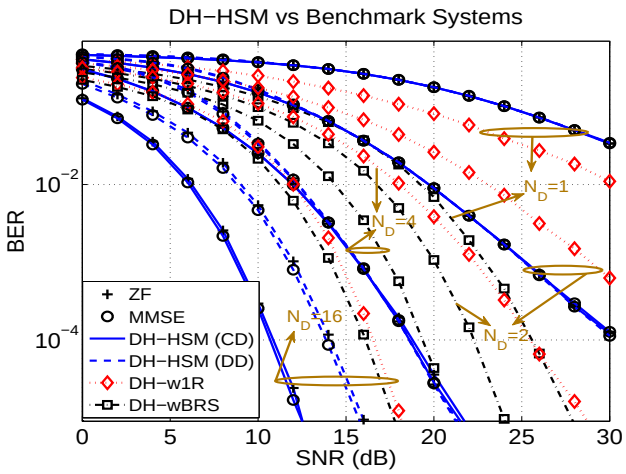
6.5.3 Medium Scale Setup: The Effect of Multiple Antennas at the DN

The performance evaluation of DH-HSSK/HSM is presented in Fig. 6.4. The same figure shows the performance of DH-w1R and DH-wBRS for the purpose of comparison. Note that in the following results, direct link communication and distributed STBC are not used as benchmark systems. Direct link communication is not employed, because DH-HSM is proposed for a scenario where either the SN-DN is very weak (where direct link communication is clearly outperformed), or there is no SN-DN link. Also, distributed STBC is not further used, because as the size of the distributed STBC codeword increases, its fractional STBC rate determines performance significantly worse than DH-HSM.

In addition, note that for DH-w1R and DH-wBRS, the BER performance is depicted only for ZF



(a) $16 \times 8 \times N_D$ DH-HSSK system ($k_t = 1.5$ bpsp).



(b) $16 \times 8 \times N_D$ DH-HSM system ($k_t = 3.5$ bpsp).

Figure 6.4: *BER performance of a medium scale DH-HSSK/HSM system, for* $N_D = \{1, 2, 4, 16\}$*, when: i) the SN employs P-CSIT and ii) the SNR per hop is the same.*

precoding at the SN. It can be proved that in the very special case of a MISO linear precoding system (first hop) and when the normalization process of Section 6.2 is applied, the MMSE and the ZF forcing precoders result in the same receiving signal. Hence, the performance is the same.

Figs. 6.4(a) and 6.4(b) demonstrate that when the number of receive antennas is small $N_D = \{1, 2\}$, there is no performance difference between the CD and DD algorithms of DH-HSSK and DH-HSM. As N_D is increased, there is a difference. For $N_D = 4$, the difference can be noticed only in low SNR. However, when $N_D = 16$, this performance gap is significantly increased. This performance difference exists due to the lost multiplexing gain which is caused

by the activation of multiple RNs.

Regarding the comparison with the benchmark systems, it can be seen that for $N_D > 1$, DH-HSSK and DH-HSM are significantly better than DH-w1R. When, $N_D = 1$, DH-HSSK offer a sightly better performance than DH-w1R. Also, DH-HSM is outperformed due to the unity diversity order of SM during the second hop [143].

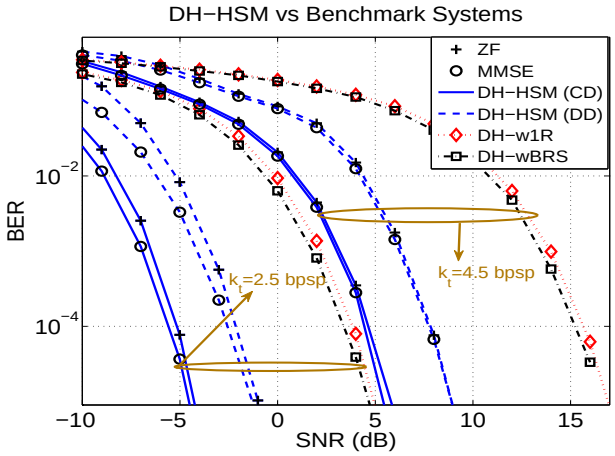
In addition, Figs. 6.4(a) and 6.4(b) include the comparison of DH-HSSK and DH-HSM with DH-wBRS. Recall that this comparison is not absolutely fair for DH-HSSK and DH-HSM. For $N_D = \{1, 2\}$, DH-wBRS demonstrates better performance. In contrast, for a value of $N_D = 4$ and low SNR, DH-HSSK and DH-HSM outperform DH-wBRS due to their higher multiplexing gain. However, after a certain point (~ 7.5 dB and ~ 20 dB for DH-HSSK and DH-HSM, respectively), DH-wBRS has a better performance due to its higher diversity order. Remarkably, when $N_D = 16$ and for the depicted practical BER, DH-HSSK and DH-HSM offer a significantly better performance, even for DD.

6.5.4 A Very Large Scale Setup

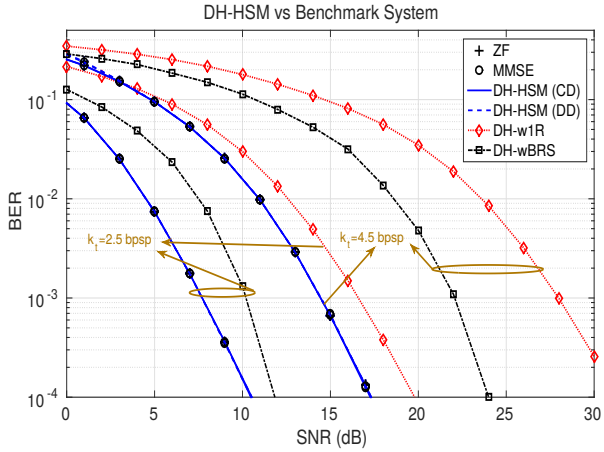
Recently, the concept of very large scale MIMO (or Massive MIMO) has attracted great attention from the research community [86]. Thus, Fig. 6.5 demonstrates the performance of DH-HSSK and DH-HSM in a system with a large number of antennas at the RN and DN, and a large number of RNs ($64 \times 32 \times N_r$, where, $N_r = \{64, 4\}$).

The first conclusion drawn from Fig. 6.5(a) is that both DH-HHSK and DH-HSM achieve practical BERs in very low SNR. For example, DH-HSSK achieves a BER of 10^{-4} at about -5.4 dB when CD is employed and at about -2.2 dB when DD is employed. This phenomenon is attributed to the massive MIMO setup of the first hop [86] and to the massive multiplexing gain and receive diversity obtained from SM during the second hop. In addition, Fig. 6.5(a) demonstrates that the employment of ZF or MMSE precoding at the SN results in a performance gap for both DH-HSSK and DH-HSM architectures, and for both CD and DD. Regarding the use of CD or DD at the RNs, it can be observed that there is a penalty for both DH-HSSK and DH-HSM when DD is employed. This penalty is decreased as the value of SNR is decreased.

Furthermore, Fig. 6.5(a) shows that DH-HSSK ($k_t = 2.5$ bpsp) and DH-HSM ($k_t = 4.5$ bpsp), both for CD and DD, have a significantly better performance than DH-w1R and DH-wBRS, for the depicted practical values of BER. The reader is reminded that the comparison between



(a) $N_r = 64$



(b) $N_r = 4$

Figure 6.5: *BER performance of a very large scale DH-HSSK/HSM* ($64 \times 32 \times N_r$*, where,* $N_r = \{64, 4\}$*) system when: i) the SN employs P-CSIT, and ii) the SNR per hop is the same.*

DH-HSSK, DH-HSM, and DH-wBRS is not entirely fair, due to the closed loop transmission of the second hop of DH-wBRS. It is quite notable, that for BER=10^{-4}, DH-HSSK has a performance difference of about 8 dB when CD is employed and 5 dB when DD is employed at the RNs, compared with DH-wBRS. The performance difference against DH-w1R is further increased to about 9 dB when CD is employed and 6 dB when DD is employed. Additionally, for the same BER, DH-HSM demonstrates a performance gap of about 9 dB for CD and about 6 dB for DD, compared with DH-wBRS. However, when DH-HSM is compared with DH-w1R, the performance difference is further increased to about 10.5 dB for CD and to 7.5 dB for DD. Note that all systems are able to achieve a large receive diversity gain due to the big number of

Configuration		k (bpsp)	Target BER	RAER [%]
DH-HSSK	$8 \times 4 \times 1$	1	10^{-2}	0%
with CD	$16 \times 8 \times 4$	1.5	10^{-3}	77.6%
at the RNs	$64 \times 32 \times 64$	2.5	10^{-4}	87%
DH-HSSK	$8 \times 4 \times 1$	1	10^{-2}	0%
with DD	$16 \times 8 \times 4$	1.5	10^{-3}	77.6%
at the RNs	$64 \times 32 \times 64$	2.5	10^{-4}	75%
DH-HSM	$8 \times 4 \times 1$	3	10^{-1}	-14.5%
with CD	$16 \times 8 \times 4$	3.5	10^{-3}	77.6%
at the RNs	$64 \times 32 \times 64$	4.5	10^{-4}	91%
DH-HSM	$8 \times 4 \times 1$	3	10^{-1}	-14.5%
with DD	$16 \times 8 \times 4$	3.5	10^{-3}	77.6%
at the RNs	$64 \times 32 \times 64$	4.5	10^{-4}	82.2%

Table 6.1: *Relative energy efficiency gains of DH-HSSK/HSM versus DH-w1R for different system configurations.*

receive antennas at the DN.

Similar conclusions as Fig. 6.5(a) can be drawn for the system setup presented in Fig. 6.5(b). However, there are two differences. Firstly, both for DH-HSSK and DH-HSM, and for the depicted SNR, there is no performance difference between CD and DD. And secondly, all schemes no further achieve a BER performance of interest in low SNR as Fig. 6.5(a). In particular, since the DN deploys only two receive antennas, there is no massive multiplexing gain and receive diversity during the second hop.

At this point, it is emphasized that due to the simplicity, low complexity, and very good BER performance of DH-HSSK and DH-HSM at very low SNR, the new architecture would be a perfect candidate for wireless backhaul connection between BSs with a very large number of antennas. Such an application could be easily implemented using either fixed or mobile RNs, or a combination of them.

6.5.5 RF Energy Efficiency Gains of DH-HSSK/HSM

In order to quantify the RF energy efficiency gain of DH-HSSK and DH-HSM over the benchmark system of DH-w1R, the metric of the RAER is deployed. This metric is defined in (4.92). Table 6.1 demonstrates the RAER benefits of DH-HSSK and DH-HSM over DH-w1R, for different system configurations and spectral efficiency (in bpsp). In order to conduct a fair comparison between DH-HSSK, DH-HSM, and the benchmark system, the same BER target and spectral efficiency is used for each system configuration. As it can be seen from Table 6.1, in almost all studied scenarios, DH-HSSK and DH-HSM are more energy efficient than DH-w1R.

There are only two exceptions, when the DN is equipped with one receive antenna ($N_D = 1$). In all other scenarios, DH-HSSK and DH-HSM achieve significant energy efficiency gains (in terms of RAER) over DH-w1R. As it can be seen, these gains can be as high as 91%.

Finally, note that the computational complexity analysis of the first hop of DH-HSM can be directly obtained from Section 4.7.6, as the first hop of DH-HSM corresponds to R-SM transmission. Also, the computational complexity analysis of the second hop of DH-HSM is provided in the discussion of (6.7). The discussion regarding the effect of the number of RNs in the system complexity is presented in detail in the introduction of Section 6.5. In this way, by taking into account the computational complexity of the new architecture, its BER performance and EE benefits, it can be concluded that it offers the best trade-off between EE, complexity, and data rate whenever it offers better BER performance than the benchmark systems.

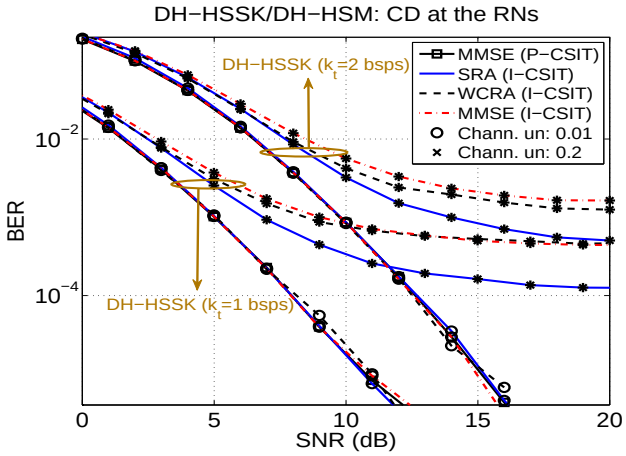
6.5.6 BER performance of DH-HSSK/HSM: The Effect of I-CSIT at the SN

In order to acquire a complete picture of the performance of DH-HSSK/HSM under: i) the effect of I-CSIT and ii) different types of precoding (MMSE, SRA, and WCRA) at the SN, Fig. 6.6 presents the BER curves for both systems when the variance of the SN-RNs channel uncertainty is $\sigma^2_{\widetilde{\mathbf{H}}_{\text{SR}}} = \{0.01, 0.2\}$. For the sake of comparison, in the same figures, the performance of DH-HSSK and DH-HSM is included when MMSE precoding with P-CSIT is employed at the SN. In addition, is is assumed that the SN-RNs and RNs-DN channels are uncorrelated[2].

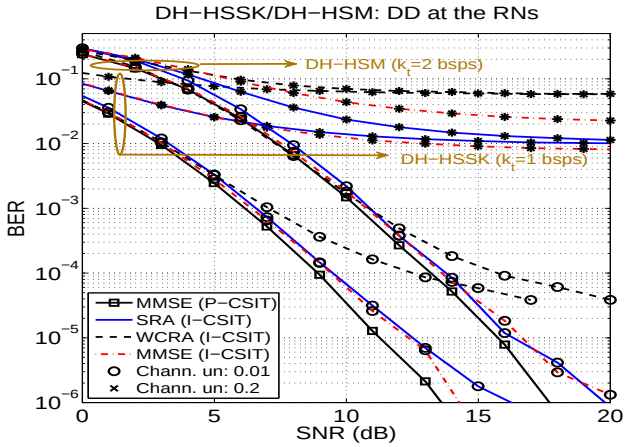
As it can be seen from Fig. 6.6(a), when CD is employed at the RNs and the channel uncertainty is small ($\sigma^2_{\widetilde{\mathbf{H}}_{\text{SR}}} = 0.01$), DH-HSSK and DH-HSM have no performance difference from the ideal scenario of P-CSIT at the SN. This is true for all types of precoding. In contrast, the increase of channel uncertainty to $\sigma^2_{\widetilde{\mathbf{H}}_{\text{SR}}} = 0.2$ causes a BER performance degradation. In low SNR, the degradation is about 1 dB. In addition, an error floor is caused in high SNR, due to the noise limited detection at the RNs. In fact, this is the effect of inaccurate beamforming from the SN. Finally, from Fig. 6.6(a), it is concluded that when I-CSIT is available at the SN, the optimal precoding method is SRA.

In addition, Fig. 6.6(b) demonstrates that the effect of I-CSIT has a more diminishing result for the case of DD. In fact, in low SNR, there is a small BER performance loss even for low

[2]Note that, in order to provide a fair comparison between the different precoding methods, the normalization matrix $\mathbf{D}$ is used as defined in Section 6.2.



(a) CD at the RNs.

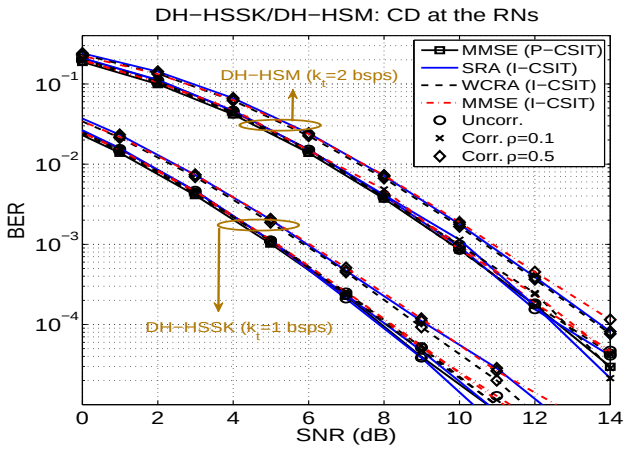


(b) DD at the RNs.

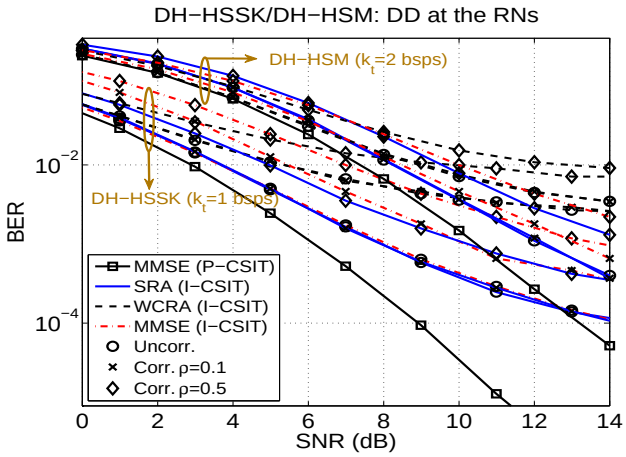
Figure 6.6: *BER performance of a* $8 \times 4 \times 4$ *DH-HSSK* ($k_t = 1$ *bpsp) and a* $8 \times 4 \times 4$ *DH-HSM* ($k_t = 2$ *bpsp) system, when different values of channel uncertainty for the SN-RNs channel* ($\sigma^2_{\widetilde{\mathbf{H}}_{SR}} = \{0.01, 0.2\}$) *are used.*

channel uncertainty ($\sigma^2_{\widetilde{\mathbf{H}}_{SR}} = 0.01$), both for DH-HSSK and DH-HSM. Also, after 10 dB, the phenomenon of the error floor in high SNR starts to appear. Finally, it can be seen that high channel uncertainty ($\sigma^2_{\widetilde{\mathbf{H}}_{SR}} = 0.2$) causes a diminishing BER performance, even in high SNRs.

In Fig. 6.7, the BER performance of DH-HSSK/DH-HSM is demonstrated under the composite effect of: i) I-CSIT at the SN and ii) correlated SN-RNs and RNs-DN channel. The variance of channel uncertainty at the SN is $\sigma^2_{\widetilde{\mathbf{H}}_{SR}} = 0.05$. Furthermore, the considered correlation scenarios are: i) low correlation with $\rho = 0.1$ and ii) high correlation with $\rho = 0.5$. In addition, given that the RNs are placed far apart, the SN-RNs channel contains only transmit space correlation, while the RNs-DN channel is affected only by receive space correlation. As



(a) CD at the RNs.



(b) DD at the RNs.

Figure 6.7: *BER performance of a* $8 \times 4 \times 4$ *DH-HSSK* ($k = 1$ *bpsp) and a* $8 \times 4 \times 4$ *DH-HSM* ($k = 2$ *bsps) system when: i) both the SN-RN and the RN-DN channels are correlated with* $\rho = \{0.1, 0.5\}$ *and ii) the SN possesses I-CSIT with* $\sigma^2_{\widetilde{\mathbf{H}}_{SR}} = 0.05$.

a benchmark reference, the ideal scenario of P-CSIT at the SN and uncorrelated SN-RNs and RNs-DN channels is deployed.

Fig. 6.7(a) shows that for the depicted SNRs and when correlation is low $\rho = 0.1$, the performance of the proposed architecture with CD does not deviate from the ideal case of no correlation and P-CSIT at the SN. Furthermore, the observation of the same figure shows that high correlation degrades the BER performance for about 1 dB at BER=10^{-3}. In fact, the previous conclusions are practically true for all methods of precoding. Also, Fig. 6.7(b) presents that the use of DD at the RNs results in a diminishing behavior, even for low correlation. Finally, Fig. 6.7(b) demonstrates that SRA is the best precoding method under the evaluated system setup.

6.6 Summary

The main objective of Chapter 6 was the extension of the concept of space modulation to a VMIMO architecture. Based on the concepts of R-SM and SM, an architecture which transfers information using the spatial position of multiple RNs, in addition, or not, to the conventional way was formed. Using MIMO linear precoding, the native problem of the activation of multiple RNs of the space modulated VMIMO schemes was reduced or practically avoided.

The new architecture was compared against the appropriate SotA schemes and demonstrated significant BER gains. Especially, for a very large system setup and BER=10^{-4}, the performance difference was up to 9 dB for DH-wBRS and up to 10 dB for DH-w1R. Hence, it was concluded that the concept of space modulation can be successfully applied in a distributed framework. Also, strict bounds for the ABEP of the new architecture were provided, both for the CD and DD algorithms at the RNs. Finally, the effects of I-CSIT at the SN and channel correlation were studied. Using simulation results, it was concluded that the new architecture and precoding methods are robust to realistic scenarios of CSIT and correlation.

Chapter 7
Conclusions, Limitations, and Future Work

7.1 Summary and Conclusions

In this thesis, the challenges of the energy evaluation of SM and its extension in new MIMO setups were addressed. More specifically, the EARTH power model was deployed in order to identify the EE gains of SM in a BS when TDMA is the multiple access scheme. In addition, the performances of R-SM and MSR-SM were studied using a wide range of theoretical and simulation results. Both R-SM and MSR-SM are based on the incorporation of the concept of space modulation at the receiver side. Furthermore, a form of the broadcast channel which deploys R-SM or MSR-SM was proposed and studied. Finally, SM and R-SM were combined in order to form a novel and very efficient dual hop VMIMO architecture. In the following, a summary of the main contributions and conclusions of this thesis is provided.

In Chapter 3, the theoretical limit of channel capacity and Monte Carlo simulation results were deployed in order to display the EE gains of SM in different types of BS and when TDMA is used. The types of BS considered are: i) macro; ii) micro; iii) pico; and iv) femtocell BS. It was concluded that the single RF chain configuration of SM results in high EE gains over the corresponding conventional MIMO techniques. In this context, it was shown that SM can be up to 67% more energy efficient than MISO with only transmit diversity [6], STBC [29], and MIMO [30], under all types of BSs. Furthermore, it was inferred that SM and the deployed benchmark systems, for all types of BS, achieve a transmission point where the EE (in Mbits/J) is maximized. In addition, it was concluded that for a range of data rates, SM provides the same data rate as the benchmark systems with significantly less power supply. For example, for four transmit antennas, SM has a reduction which can be as low as 76% compared to the corresponding conventional MIMO schemes.

In Chapter 4, the focus was on the evaluation of the performance of R-SM and MSR-SM using analytical and Monte Carlo simulation results. More specifically, the theoretical ABEP of

R-SM and MSR-SM was provided, using the union bound technique, when: i) the wireless channel follows the Rayleigh distribution; ii) the transmitter possesses P-CSIT; and iii) the deployed precoder is based on the ZF principle. Note that the characterization of the received signal was done using a new accurate statistical framework. Therefore, it was concluded that the new bounds are accurate. Furthermore, the diversity order and coding gain of R-SM and MSR-SM were derived. It was concluded that MSR-SM with ZF precoding achieves the same diversity order and higher coding gain as the corresponding conventional MIMO benchmark system. The performances of R-SM and MSR-SM were compared against the one of the corresponding conventional MIMO. It was inferred that R-SM offers the same or better BER performance than conventional MIMO only when the spectral efficiency is low and the number of transmit and receive antennas is low. Thus, R-SM cannot be considered as a suitable candidate for a massive MIMO setup. Also, it was concluded that when the number of the multiple parallel spatially modulated data streams is close to the number of receive antennas, MSR-SM has a better BER performance than the corresponding conventional MIMO. Furthermore, two regularized linear precoding methods, which deploy realistic I-CSIT, were proposed. The first precoder deploys I-CSIT in a statistical form. In contrast, the second precoding design uses I-CSIT in a worst case form. Using analytical arguments, the instantaneous and the average power of all precoders used in this thesis were derived. The new precoders were shown to have a good performance BER performance under different channel correlation states.

In Chapter 5, the concept of receive space modulation was deployed in order to form a SDMA architecture for the downlink. Under the assumptions of P-CSIT and ZF precoding, upper bounds for the ABEP of the new scheme were provided. The new bounds take into account the large scale effect. Using Monte Carlo simulation results, it was deduced that the new bounds are tight in high SNR. Furthermore, it was inferred that the diversity order of the new architecture is not affected by the large scale effect, when the large scale effect is deterministic. The performance of the new architecture was compared against the one of corresponding conventional MIMO scheme. It was shown that the new scheme is able to have a better BER performance in high SNR.

The incorporation of SM and R-SM in a VMIMO architecture was presented in Chapter 6. The new scheme is termed DH-HSM. In the new architecture, multiple RNs cooperate in order to establish communication between a SN and a DN. This is achieved by spatially modulating the RNs. It was revealed that, using MIMO linear precoding, DH-HSM practically avoids

the native problem of the activation of multiple RNs. Furthermore, using simulation results, it was demonstrated that DH-HSM offers significant BER gains against the appropriate SotA schemes. In more detail, the performance difference was up to 9 dB for DH-wBRS and up to 10 dB for DH-w1R, when a very large system setup and BER=10^{-4} were considered. Hence, it was concluded that the concept of DH-HSM can be successfully applied in very large system setups. In addition, new strict bounds for the ABEP of DH-HSM were derived. Finally, the performance of DH-HSM was studied under the effects of I-CSIT at the SN and channel correlation. Using simulation results, it was concluded that the new architecture and new precoding methods are robust to realistic scenarios of CSIT and correlation.

7.2 Limitations and Future Work

Although this thesis conducted a detailed study on the evaluation and extension of SM in different system setups, there are some aspects that would be worthwhile to be researched in the future. In Chapter 3, the evaluation of the EE of SM in a BS was conducted when the multiple access scheme is TDMA. Due to the popularity of OFDMA in modern cellular systems, the energy evaluation of SM needs to be extended in this multiple access scheme. The incorporation of SM in a cellular system which deploys OFDMA can be done by using the same principles as any other MIMO transmission scheme [188]. In particular, provided that each transmit antenna is equipped with an OFDM modulator and each one of the multiple users is allocated a portion of the available resources (in time and frequency domain), the formation of an OFDMA based spatially modulated system can be directly obtained from [14]. Such a scheme requires a number of RF chains equal to the number of transmit antennas. Therefore, the EE benefits of SM which are obtained due to the single RF chain configuration vanish. However, in the frequency domain, the multiple SM transmitted signals may result in a sparse signal transmission. Actually, the sparsity of the transmitted signal of SM promises better Peak-to-Average-Power Ratio (PAPR) than conventional MIMO techniques. Therefore, some of the EE gains of SM may be retained in an OFDMA configuration. Furthermore, Chapter 3 provided only simulation results. Therefore, an analytical confirmation of these results would be an interesting subject of research.

A detailed evaluation of R-SM and MSR-SM in single-user and multi-user setup was presented in Chapters 4 and 5, respectively. These studies were conducted using simulation and theoretical results under the assumptions of P-CSIT and ZF precoding. For all other types of precoding

and CSIT, only simulation results were presented. Although a theoretical study of the performance of R-SM and MSR-SM under I-CSIT and under other methods of precoding is challenging, such a study would provide a better insight to the performance of the new architectures. In addition, the development of such a theoretical framework could be readily incorporated in conventional MIMO schemes.

A VMIMO scheme based on the concepts of SM and R-SM was proposed in Chapter 6. Although Chapter 6 includes a wide range of simulation and analytical results, the theoretical evaluation of the ABEP of DH-HSM when DD is deployed at the RNs was undertaken semi-analytically. In fact, this study is semi-analytical due to the challenging derivation of the PDF of the RV of (6.37). Therefore, this derivation would be important. In addition, an interesting research study would be the evaluation of the performance of DH-HSM when the RNs deploy the AF protocol. Finally, the incorporation of MSR-SM and MS-SM in a setup similar to DH-HSM may exhibit performance gains and would be a worthwhile subject of further research.

Appendices

Appendix A
Evaluation of the Integral of (4.84)

If the marginal PDF of λ (given from (4.83)) is incorporated in (4.84), the integral that needs to be evaluated is transformed to:

$$\mathrm{E}_{\lambda}\left[\frac{\lambda}{(\lambda+\phi)^2}\right]=\frac{1}{N_r}\sum_{k=0}^{N_r-1}\int_0^{+\infty}\frac{x^{\Delta+1}}{(x+\phi)^2}e^{-x}\frac{k!}{(k+\Delta)!}\left[L_k^{\Delta}(x)\right]^2\mathrm{d}x. \tag{A.1}$$

Provided that:

$$\left[L_k^{\Delta}(x)\right]^2=\sum_{l=0}^{k}\sum_{m=0}^{k}\frac{(-1)^{l+m}}{l!m!}\binom{k+\Delta}{k-l}\binom{k+\Delta}{k-m}x^{l+m}, \tag{A.2}$$

as the square of a Laguerre polynomial (square of a polynomial), after some straightforward arithmetic operations, (A.1) can be transformed to:

$$\begin{aligned}\mathrm{E}_{\lambda}\left[\frac{\lambda}{(\lambda+\phi)^2}\right]=\frac{1}{N_r}\sum_{k=0}^{N_r-1}\frac{k!}{(k+\Delta)}\sum_{l=0}^{k}\sum_{m=0}^{k}\frac{(-1)^{l+m}}{l!m!}\binom{k+\Delta}{k-l}\binom{k+\Delta}{k-m}\\ \times I_A(l+m+\Delta+1,\phi),\end{aligned} \tag{A.3}$$

where,

$$I_A(n,\phi)=\int_0^{+\infty}x^n e^{-x}\frac{1}{(x+\phi)^2}\mathrm{d}x. \tag{A.4}$$

In the next step, the method of integration by parts is applied in (A.4). In this way, it is shown that:

$$\begin{aligned}I_A(n,\phi)&=\int_0^{+\infty}x^n e^{-x}\frac{1}{(x+\phi)^2}\mathrm{d}x\\ &=\int_0^{+\infty}x^n e^{-x}\left(\frac{-1}{x+\phi}\right)'\mathrm{d}x\\ &=\frac{-x^n e^{-x}}{x+\phi}\Big|_0^{+\infty}+n\int_0^{+\infty}\frac{x^{n-1}e^{-x}}{x+\phi}\mathrm{d}x-\int_0^{+\infty}\frac{x^n e^{-x}}{x+\phi}\mathrm{d}x.\end{aligned} \tag{A.5}$$

It is not difficult to see that:

$$\left.\frac{-x^n e^{-x}}{x+\phi}\right|_0^{+\infty} = \lim_{x\to\infty} \frac{-x^n e^{-x}}{x+\phi} - 0 = 0. \tag{A.6}$$

In addition, from [170, p. 348], it holds that:

$$\int_0^{+\infty} \frac{x^{n-1}e^{-x}}{x+\phi}\mathrm{d}x = \phi^{n-1}e^{\phi}\Gamma(n)\Gamma(1-n,\phi). \tag{A.7}$$

Thus, if the two integrals of (A.5) are evaluated using (A.7) and the result of (A.6) is taken into account, after some manipulations, it can be shown that:

$$I_A(n,\phi) = n!\phi^{n-1}e^{\phi}\left[\Gamma(1-n,\phi) - \phi\Gamma(-n,\phi)\right]. \tag{A.8}$$

In the final step of this derivation, (A.8) is incorporated in (A.9), which gives that:

$$\begin{aligned}\mathrm{E}_\lambda\left[\frac{\lambda}{(\lambda+\phi)^2}\right] = \frac{1}{N_r}\sum_{k=0}^{N_r-1}\frac{k!}{(k+\Delta)}\sum_{l=0}^{k}\sum_{m=0}^{k}\frac{(-1)^{l+m}}{l!m!}\binom{k+\Delta}{k-l}\binom{k+\Delta}{k-m}\\ \times (l+m+\Delta+1)!\phi^{l+m+\Delta}e^{\phi}\\ \times\left[\Gamma(-l-m-\Delta,\phi) - \phi\Gamma(-l-m-\Delta-1,\phi)\right].\end{aligned} \tag{A.9}$$

Appendix B
Evaluation of the Integral of (4.88)

In this appendix, the evaluation of the expectation of (4.88) is provided. This expectation is given as:

$$\mathrm{E}_{\bar{\lambda}}\left[\frac{\bar{\lambda}}{(\bar{\lambda}+\vartheta)^2}\right]=\int_0^{+\infty}\frac{x}{(x+\vartheta)^2}f_{\bar{\lambda}}(x)\mathrm{d}x. \tag{B.1}$$

Using the marginal PDF of $\bar{\lambda}$, which is available in (4.87), after some manipulations (B.1) is re-written as it follows:

$$\begin{aligned}\mathrm{E}_{\bar{\lambda}}\left[\frac{\bar{\lambda}}{(\bar{\lambda}+\vartheta)^2}\right]&=K\int_0^{+\infty}\sum_{i=1}^{N_r}\sum_{j=1}^{N_r}\frac{x^j}{(x+\vartheta)^2}\mathcal{D}(i,j)\left(\alpha_{\varrho+i}^{\varrho-1}e^{-\frac{x}{\alpha_{\varrho+i}}}\right.\\&\left.-\sum_{l=1}^{\varrho}\sum_{k=1}^{\varrho}\left[\boldsymbol{\Psi}^{-1}\right]_{k,l}\alpha_{\varrho+i}^{k-1}\alpha_l^{\varrho-1}e^{-\frac{x}{\alpha_l}}\right)\mathrm{d}x\\&=KH,\end{aligned} \tag{B.2}$$

where,

$$H=\int_0^{+\infty}\sum_{i=1}^{N_r}\sum_{j=1}^{N_r}\frac{x^j}{(x+\vartheta)^2}\mathcal{D}(i,j)\left(\alpha_{\varrho+i}^{\varrho-1}e^{-\frac{x}{\alpha_{\varrho+i}}}-\sum_{l=1}^{\varrho}\sum_{k=1}^{\varrho}\left[\boldsymbol{\Psi}^{-1}\right]_{k,l}\alpha_{\varrho+i}^{k-1}\alpha_l^{\varrho-1}e^{-\frac{x}{\alpha_l}}\right)\mathrm{d}x. \tag{B.3}$$

After a straightforward elaboration on (B.3), it can be shown that H can be written as:

$$\begin{aligned}H=\sum_{i=1}^{N_r}\sum_{j=1}^{N_r}\mathcal{D}(i,j)\left[\alpha_{\varrho+i}^{\varrho-1}\left(\int_0^{+\infty}\frac{x^j}{(x+\vartheta)^2}e^{-\frac{x}{\alpha_{\varrho+i}}}\mathrm{d}x\right)\right.\\\left.-\sum_{l=1}^{\varrho}\sum_{k=1}^{\varrho}\left[\boldsymbol{\Psi}^{-1}\right]_{k,l}\alpha_{\varrho+i}^{k-1}\alpha_l^{\varrho-1}\left(\int_0^{+\infty}\frac{x^j}{(x+\vartheta)^2}e^{-\frac{x}{\alpha_l}}\mathrm{d}x\right)\right]\end{aligned} \tag{B.4}$$

As it can be seen from (B.4), two integrals which have the following form:

$$I_B(\vartheta,\nu)=\int_0^{+\infty}\frac{x^j}{(x+\vartheta)^2}e^{-\nu x}\mathrm{d}x, \tag{B.5}$$

have to be evaluated. In order to deal with (B.5), the method of integration by parts is applied.

In this way, it is obtained that:

$$I_B(\vartheta,\nu)=\int_0^{+\infty}x^j e^{-\nu x}\left(-\frac{1}{x+\vartheta}\right)'\mathrm{d}x=\frac{-x^j e^{-\nu x}}{x+\theta}\bigg|_0^{+\infty}+\int_0^{+\infty}\left(x^j e^{-\nu x}\right)'\frac{1}{x+\vartheta}\mathrm{d}x$$
$$=j\int_0^{+\infty}\frac{x^{j-1}e^{-\nu x}}{x+\vartheta}\mathrm{d}x-\nu\int_0^{+\infty}\frac{x^j e^{-\nu x}}{x+\vartheta}\mathrm{d}x. \tag{B.6}$$

In (B.6), it is used the fact that:

$$\frac{-x^j e^{-\nu x}}{x+\theta}\bigg|_0^{+\infty}=\lim_{x\to+\infty}\frac{-x^j e^{-\nu x}}{x+\theta}-0=0. \tag{B.7}$$

By evaluating the two integrals of (B.6), using (A.7) which is a direct result from [170, p. 348], it can be proven that:

$$I_B(\vartheta,\nu)=j\vartheta^{j-1}e^{\vartheta\nu}\Gamma(j)\Gamma(1-j,\vartheta\nu)-\nu\vartheta^j e^{\vartheta\nu}\Gamma(j+1)\Gamma(1-j,\vartheta\nu). \tag{B.8}$$

In this way, (B.8) is plugged in (B.4) which gives:

$$H=\sum_{i=1}^{N_r}\sum_{j=1}^{N_r}\mathcal{D}(i,j)\left[\alpha_{\varrho-i}^{\varrho-1}g_j\left(\frac{1}{\alpha_{\varrho+i}}\right)-\sum_{l=1}^{\varrho}\sum_{k=1}^{\varrho}\left[\mathbf{\Psi}^{-1}\right]_{k,l}\alpha_{\varrho+i}^{k-1}\alpha_l^{\varrho-1}g_j\left(\frac{1}{\alpha_l}\right)\right], \tag{B.9}$$

where,

$$g_j(y)=\vartheta^{j-1}e^{\vartheta y}j!\left[\Gamma(1-j,y\vartheta)-y\vartheta\Gamma(-j,y\vartheta)\right]. \tag{B.10}$$

Thus, the incorporation of (B.9) in (B.2) results in:

$$\mathrm{E}_{\bar{\lambda}}\left[\frac{\bar{\lambda}}{(\bar{\lambda}+\vartheta)^2}\right]=K\sum_{i=1}^{N_r}\sum_{j=1}^{N_r}\mathcal{D}(i,j)\left[\alpha_{\varrho+i}^{\varrho-1}g_j\left(\frac{1}{\alpha_{\varrho+i}}\right)-\sum_{l=1}^{\varrho}\sum_{k=1}^{\varrho}\left[\mathbf{\Psi}^{-1}\right]_{k,l}\alpha_{\varrho+i}^{k-1}\alpha_l^{\varrho-1}g_j\left(\frac{1}{\alpha_l}\right)\right]. \tag{B.11}$$

In this way, the derivation is completed.

Appendix C
Evaluation of the Expectation in (5.43)

This appendix provides the evaluation of the expectation in (5.43). Let the PDF of z_i in (5.38), (5.43) can be expressed as follows:

$$\begin{aligned} g\left(y\right) &= \mathrm{E}_{z_i}\left[e^{-y\breve{\gamma}_i z_i}\right] = \int_{-\infty}^{+\infty} e^{-y\breve{\gamma}_i x} f_{z_i}(x)\mathrm{d}x \\ &= \left[\prod_{l=1}^{N_i}\left(\frac{\breve{\alpha}_1}{\breve{\alpha}_l}\right)^{L_{\mathrm{MU}}}\right]\left[\int_0^{+\infty} e^{-y\breve{\gamma}_i x}\right. \\ &\times \left.\sum_{k=0}^{+\infty}\frac{\breve{\delta}_k x^{N_i L_{\mathrm{MU}}+k-1} e^{-\frac{x}{\breve{\alpha}_1}}}{\breve{\alpha}_1^{N_i L_{\mathrm{MU}}+k}\Gamma(N_i L_{\mathrm{MU}}+k)}\mathrm{d}x\right] \\ &= \left[\prod_{l=1}^{N_i}\left(\frac{\breve{\alpha}_1}{\breve{\alpha}_l}\right)^{L_{\mathrm{MU}}}\right]\sum_{k=0}^{+\infty}\left[\frac{\breve{\delta}_k}{\breve{\alpha}_1^{N_i L_{\mathrm{MU}}+k}\Gamma(N_i L_{\mathrm{MU}}+k)}\right. \\ &\times \left.\int_0^{+\infty} x^{N_i L_{\mathrm{MU}}1+k-1} e^{-\left(y\breve{\gamma}_i+\frac{1}{\breve{\alpha}_1}\right)x}\mathrm{d}x\right], \end{aligned} \tag{C.1}$$

From the integration formula in [170, p.346, 3.381, 4], it holds that:

$$\int_0^{+\infty} x^{\nu-1} e^{-\mu x}\mathrm{d}x = \mu^{-\nu}\Gamma(\nu), \tag{C.2}$$

where $\nu > 0$ and $\mathrm{Re}\{\mu\} > 0$, $\Gamma(\cdot)$ denotes the incomplete gamma function defined in [170, p. 899].

By plugging (C.2) in (C.1) and with the aid of some manipulations, it is given that:

$$\begin{aligned} g\left(y\right) &= \left[\prod_{l=1}^{N_i}\left(\frac{\breve{\alpha}_1}{\breve{\alpha}_l}\right)^{L_{\mathrm{MU}}}\right]\left(y\breve{\alpha}_1\breve{\gamma}_i+1\right)^{-N_i L_{\mathrm{MU}}} \\ &\times \sum_{k=0}^{+\infty}\breve{\delta}_k\left(y\breve{\alpha}_1\breve{\gamma}_i+1\right)^{-k}, \end{aligned} \tag{C.3}$$

from which the proof follows.

Appendix D
High SNR Approximation of the PEP of (5.44)

If the SNR approaches to infinity ($\gamma \to +\infty$), only the smallest value of the exponent k in (5.47) needs to be considered, which is equal to one. Thus, (5.47) can be approximated as:

$$\begin{aligned} P_i^{+\infty}(\mathbf{x}_i \to \hat{\mathbf{x}}_i, \gamma) &\approx \frac{\left[\prod_{l=1}^{N_i} \left(\frac{\breve{\alpha}_1}{\breve{\alpha}_l}\right)^{L_{\mathrm{MU}}}\right]}{2} \left(\frac{\breve{\alpha}_1}{4}\xi_i\gamma + 1\right)^{-N_i L_{\mathrm{MU}}} \\ &= \left[\sqrt[N_i L_{\mathrm{MU}}]{\frac{2}{\prod_{l=1}^{N_i} \left(\frac{\breve{\alpha}_1}{\breve{\alpha}_l}\right)^{L_{\mathrm{MU}}}}} \frac{\breve{\alpha}_1}{4}\xi_i\gamma + \sqrt[N_i L_{\mathrm{MU}}]{\frac{2}{\prod_{l=1}^{N_i} \left(\frac{\breve{\alpha}_1}{\breve{\alpha}_l}\right)^{L_{\mathrm{MU}}}}} \right]^{-N_i L_{\mathrm{MU}}} \\ &\approx \gamma^{-N_i L_{\mathrm{MU}}} \left[\frac{\prod_{l=1}^{N_i} \left(\frac{\breve{\alpha}_1}{\breve{\alpha}_l}\right)^{L_{\mathrm{MU}}}}{2} \left(\frac{\breve{\alpha}_1}{4}\xi_i\right)^{-N_i L_{\mathrm{MU}}} \right] + \mathrm{o}\left(\gamma^{-N_i L_{\mathrm{MU}}}\right) \\ &\leq \left[\frac{\breve{\alpha}_1 \sqrt[N_i L_{\mathrm{MU}}]{2}}{4}\xi_i\gamma \right]^{-N_i L_{\mathrm{MU}}} + \mathrm{o}\left(\gamma^{-N_i L_{\mathrm{MU}}}\right). \end{aligned} \tag{D.1}$$

The last step in (D.1) follows from the inequality:

$$\prod_{l=1}^{N_i} \left(\frac{\breve{\alpha}_1}{\breve{\alpha}_l}\right)^{L_{\mathrm{MU}}} \leq 1, \tag{D.2}$$

which holds because $\breve{\alpha}_l$, $l = 1, \ldots, N_i$, are the eigenvalues of $\breve{\mathbf{A}}$ in (5.39) in ascending order. This concludes the proof.

Appendix E
List of Publications

This appendix presents a list of all published and accepted papers:

1. **A. Stavridis**, D. Basnayaka, S. Sinanovic, M. Di Renzo, and H. Haas, "A Virtual MIMO Dual-Hop Architecture Based on Hybrid Spatial Modulation," *IEEE Transactions on Communications*, vol.62, no.9, pp.3161,3179, Sept. 2014.

2. **A. Stavridis**, M. Di Renzo, and H. Haas, "Performance Analysis of Multi-Stream Receive Spatial Modulation in the MIMO Broadcast Channel," *IEEE Transactions on Wireless Communications*, October 2015 (Accepted for Publication).

3. **A. Stavridis**, S. Sinanovic, M. D. Renzo, and H. Haas,"Transmit Precoding for Receive Spatial Modulation Using Imperfect Channel Knowledge, *IEEE 75th Vehicular Technology Conference: VTC 2012-Spring (3rd GreeNet Workshop)*, May 2012, Yokohama, Japan.

4. **A. Stavridis**, S. Sinanovic, M. D. Renzo, H. Haas, and Peter Grant, "An Energy Saving Base Station Employing Spatial Modulation, *IEEE 17th Int. Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, Sept. 2012, Barcelona, Spain.

5. **A. Stavridis**, S. Sinanovic, M. D. Renzo, and H. Haas, "A Power Saving Dual-Hop Architecture Based on Hybrid Spatial Modulation, *2012 Conf. Record of the Forty Sixth Asilomar Conf. on Signals, Systems and Computers (ASILOMAR)*, Nov. 2012, Pacific Grove, CA, USA.

6. **A. Stavridis**, S. Sinanovic, M. D. Renzo, and H. Haas, "Energy Evaluation of Spatial Modulation at a Multi-Antenna Base Station, *IEEE 78th Vehicular Technology Conference: VTC 2013-Fall*, Sept. 2013. Las Vegas, USA.

7. **A. Stavridis**, S. Narayanan, M. D. Renzo, L. Alonso, H. Haas, and C. Verikoukis, "A Base Station Switching On-Off Algorithm Using Traditional MIMO and Spatial Modu-

lation, *IEEE 18th Int. Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, Sept. 2013, Berlin, Germany.

8. S. Narayanan, **A. Stavridis**, M. Di. Renzo, F. Graziosi, and H. Haas, "Distributed Spatially-Modulated Space-Time-Block-Codes, *IEEE 18th Int. Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, Sept. 2013, Berlin, Germany.

9. S. Narayanan, M. J. Chaudhary, **A. Stavridis**, M. Di. Renzo, F. Graziosi, and H. Haas, "Multi-User Spatial Modulation MIMO, *IEEE Wireless Communications and Networking Conference (WCNC) 2014*, April 2014, Istanbul, Turkey.

10. **A. Stavridis**, D. Basnayaka, S. Sinanovic, M. Di Renzo, and H. Haas, "Average Bit Error Probability of Receive-Spatial Modulation Using Zero-Forcing Precoding, *IEEE 19th Int. Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, Dec. 2014, Athens, Greece.

11. M. Ijaz, D. Tsonev, **A. Stavridis**, A. Younis, J. McKendry, E. Gu, M. D. Dawson, S. Videv, and H. Haas, "Optical Spatial Modulation OFDM using Micro LEDs", *2014 Conf. Record of the Forty Eighth Asilomar Conf. on Signals, Systems and Computers (ASILOMAR),* November 2014, Pacific Grove, CA, USA.

12. **A. Stavridis**, and H. Haas, "Performance Evaluation of Space Modulation Techniques in VLC Systems", *IEEE International Conference on Communications (ICC 2015) (1ST Visible Light Communications and Networking (VLCN) Workshop),* 8-12 June 2015, London, UK.

13. **A. Stavridis**, M. D. Renzo, and H. Haas, "On the Performance of Multi-Stream Receive Spatial Modulation in the MIMO Broadcast Channel", *2015 IEEE Global Commun. Conf. (GLOBECOM),* December 2015, San Diego, CA, USA (to appear).

Appendix F
Published Papers

This appendix contains all published and accepted, journal and conference publications [24, 73, 89, 98, 101, 137, 154, 155, 168, 179, 189–191]:

F.1 IEEE Transactions on Communications

IEEE TRANSACTIONS ON COMMUNICATIONS, VOL. 62, NO. 9, SEPTEMBER 2014 3161

A Virtual MIMO Dual-Hop Architecture Based on Hybrid Spatial Modulation

Athanasios Stavridis, *Student Member, IEEE*, Dushyantha Basnayaka, *Member, IEEE*, Sinan Sinanovic, *Member, IEEE*, Marco Di Renzo, *Senior Member, IEEE*, and Harald Haas, *Member, IEEE*

***Abstract*—In this paper, we propose a novel Virtual Multiple-Input–Multiple-Output (VMIMO) architecture based on the concept of Spatial Modulation (SM). Using a dual-hop and Decode-and-Forward protocol, we form a distributed system, called Dual-Hop Hybrid SM (DH-HSM). DH-HSM conveys information from a Source Node (SN) to a Destination Node (DN) via multiple Relay Nodes (RNs). The spatial position of the RNs is exploited for transferring information in addition to, or even without, a conventional symbol. In order to increase the performance of our architecture, while keeping the complexity of the RNs and DN low, we employ linear precoding using Channel State Information (CSI) at the SN. In this way, we form a Receive-Spatial Modulation (R-SM) pattern from the SN to the RNs, which is able to employ a centralized coordinated or a distributed uncoordinated detection algorithm at the RNs. In addition, we focus on the SN and propose two regularized linear precoding methods that employ realistic Imperfect Channel State Information at the Transmitter. The power of each precoder is analyzed theoretically. Using the Bit Error Rate (BER) metric, we evaluate our architecture against the following benchmark systems: 1) single relay; 2) best relay selection; 3) distributed Space Time Block Coding (STBC) VMIMO scheme; and 4) the direct communication link. We show that DH-HSM is able to achieve significant Signal-to-Noise Ratio (SNR) gains, which can be as high as 10.5 dB for a very large scale system setup. In order to verify our simulation results, we provide an analytical framework for the evaluation of the Average Bit Error Probability (ABEP).**

***Index Terms*—MIMO, spatial modulation, linear precoding, relays, decode-and-forward.**

Manuscript received December 18, 2013; revised April 30, 2014 and July 14, 2014; accepted July 19, 2014. Date of publication July 29, 2014; date of current version September 19, 2014. This work was supported by the European Union's Seventh Program for Research, Technological Development and Demonstration under Grant PITN-GA-2010-264759. This paper was presented in part at the IEEE 75th Vehicular Technology Conference (VTC2012-Spring), Yokohama, Japan, May, 2012 and at the 46th Asilomar Conference on Signals, Systems and Computers (ASILOMAR 2012), Pacific Grove, CA, USA, November, 2012. The associate editor coordinating the review of this paper and approving it for publication was J. Yuan.

A. Stavridis, D. Basnayaka, and H. Haas are with the Institute for Digital Communications and the Joint Research Institute for Signal and Image Processing, School of Engineering, The University of Edinburgh, Edinburgh EH9 3JL, U.K. (e-mail: a.stavridis@ed.ac.uk; d.basnayaka@ed.ac.uk; h.haas@ed.ac.uk).

S. Sinanovic is with the School of Engineering and Built Environment, Glasgow Caledonian University, Glasgow G4 0BA, U.K. (e-mail: Sinan.Sinanovic@gcu.ac.uk).

M. Di Renzo is with the Laboratory of Signals and Systems (L2S), French National Center for Scientific Research (CNRS), École Supérieure d'Électricité (SUPÉLEC), University of Paris-Sud XI, 91400 Paris, France (e-mail: marco.direnzo@lss.supelec.fr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCOMM.2014.2343999

I. Introduction

OVER the recent years Multiple-Input–Multiple-Output (MIMO) communication has attracted a tremendous attention from both the academia and industry. The deployment of multi-antenna elements, at both communicating ends, is considered to be an excellent candidate solution for the future spectrally efficient wireless networks. MIMO communication offers increased data rate without the need of any further radio resources [1].

Unfortunately, MIMO comes with the major disadvantage of increased hardware complexity. Multi-antenna elements require multiple Radio Frequency (RF) chains (usually equal to the number of the employed antennas). In addition, most of the time, optimum Maximum Likelihood (ML) detection at the receiver has an exponential complexity increase with the number of receive (or transmit) antennas [2]. Other major drawbacks of MIMO are the high energy consumption due to the increased number of RF chains [3] and the complex signal processing techniques.

In order to overcome the previous native problems of MIMO, research has focused on Virtual MIMO (VMIMO) systems. The main rationale behind VMIMO is to promote low complexity system implementation by combining MIMO and relay techniques. Due to the broad nature of VMIMO, there is a wide range of designs proposed in the literature. From the MIMO point of view, there are designs that achieve diversity gains [4], multiplexing gains [5], beamforming gains [6], or a combination of them. Moreover, from the relays' point of view, there are several communication protocols. The most famous protocols are Amplify-and-Forward (AF) [6], Decode-and-Forward (DF) [4], Compress-and-Forward (CF) [7], and Estimate-and-Forward (EF) [8]. Furthermore, the previous protocols are applied in two [9] or multiple hops [4].

A. Related Work and Motivation

A new open-loop MIMO scheme that enjoys a single RF chain configuration at the transmitter and successfully treats the inherent problems of MIMO is Spatial Modulation (SM) [10]–[12]. SM is a single RF chain scheme, which is able to employ an optimum single-stream ML detector. In following we summarize the main advantages of SM: i) SM is able to achieve a multiplexing gain with the activation of a single transmit antenna; ii) SM does not require Inter-Antenna Synchronization (IAS) at the transmitter; iii) the complexity of the transmitter

does not scale significantly with the increase of antennas; and iv) SM offers increased energy efficiency. Especially, in terms of energy efficiency (bits/J), [3] show that the single RF chain configuration of SM results into significant benefits. We note that an important low-complexity and low-rate variation of SM, called Space Shift Keying (SSK) is proposed in [13]. The first real system implementation of SM is presented in [14]. Furthermore, the performance of SM under real channel measurements is performed in [15], [16]. A complete introduction on SM is provided in [17].

Until recently, MIMO precoding using Channel State Information at the Transmitter (CSIT) has been used either for the simplification of the receiver and receive power consuming benefits, or capacity gains, or the formation of Space Division Multiple Access (SDMA) schemes. In [18], MIMO precoding is used in a totally different way. Using Zero Forcing (ZF) or Minimum Mean Square Error (MMSE) precoding and the appropriate signal formation, a SM-like scheme is formed. For reasons of simplicity and in agreement with [19], we use the name of Receive-Spatial Modulation (R-SM) for this scheme. R-SM applies the space modulation concept at the receiver side.

Over the recent years, there have been many schemes that extend the concept of SM/SSK to a VMIMO distributed scenario, both for the AF and DF protocols. The authors of [20] study the performance analysis of a SSK-based AF relay network. In [21], using the idea of SSK, a cooperative transmission scheme is considered, both for the AF and DF protocols. The generalization of SSK in a two-way AF relays network is conducted in [22]. The adoption of SSK in a DF cooperative system is considered in [23]. Furthermore, the extension of SSK to a multi-branch and multi-hop scenario is studied in [24]. Moreover, the extension of Space Time Shift Keying (STSK) [25], which is a novel SM-like scheme, in a distributed VMIMO scenario is conducted in [26].

In [27], multi-antenna elements are deployed in such a way that form a dual hop scheme based on SM. The application of distributed SM to the uplink is studied in [28]. Also, [29] proposes a SSK scheme again for the uplink. In [30], using the ergodic capacity and the outage probability, it is shown that space modulation VMIMO schemes are able to achieve high throughput. Another extension of SM in a low complexity cooperative DF VMIMO architecture is conducted in [31]. The integration of Space Time Block Coding (STBC) in a distributed SM system is done in [32].

Unfortunately, all distributed and uncoded space modulation VMIMO schemes that employ multiple uncoordinated Relay Nodes (RNs) suffer from the erroneous activation of multiple RNs during the relaying phase. Thus, a Bit Error Rate (BER) performance degradation occurs. In order to overcome this problem, three strategies are followed in the literature: i) coordination between the RNs in order to ensure the activation of a single RN [33] (which is similar to the scenario of a single RN with multiple antennas [27]); ii) block-based transmission which allows the use of error correction codes at the RNs [21], [24]; and iii) advanced and error-aware detection techniques at the Destination Node (DN), which usually results in high complexity [28], [32].

B. Contributions and Outcomes

Against this background, we aim to propose and study an architecture which is suitable for the downlink. Motivated by the concept of VMIMO and the potential of SM/R-SM, we propose a non-cooperative relay architecture based on SM/R-SM. Using a half-duplex DF protocol we propose a scheme that achieves information conveyance using R-SM [19] in the first hop and SM [12] in the second hop. In this way, a Dual-Hop Hybrid Spatial Modulation (DH-HSM) system is formed. The term Hybrid Spatial Modulation (HSM) stems from the fact that our architecture employs R-SM in the first hop and SM in the second hop. The RNs are able to employ a Centralized Detection (CD) or a Distributed Detection (DD) algorithm, depending on their ability to coordinate. DH-HSM conveys information by extending the novel transmission mechanism of SM in a distributed framework.

In order to practically solve or reduce the problem of the activation of multiple RNs, without affecting the complexity of the RNs or the DN, we carefully design the transmission mode of the first hop. We use MIMO linear precoding with CSIT. In this way, we practically eliminate or reduce the effect of the wireless channel and offer significant receive Signal-to-Noise Ratio (SNR) gains at the RNs.

We consider realistic Imperfect-CSIT (I-CSIT) at the Source Node (SN) and propose two regularized linear precoding methods. The first precoder is obtained using statistical I-CSIT, while the second precoder employs worst-case I-CSIT. Furthermore, using an analytical framework, we study the instantaneous and the average power of all precoders used in this paper.

The evaluation of DH-HSM against the State-of-the-Art (SotA) is conducted using the metric of the Bit Error Rate (BER). The benchmark systems that we employ are the: i) single relay, ii) best relay selection, and iii) distributed STBC dual hop architectures, and iv) the direct communication link. Even though the comparison of DH-HSM with the best relay selection and distributed STBC is not fair for DH-HSM, our architecture is able to achieve significant BER gains in the majority of the system setups. We show that the gain of DH-HSM over the best relay selection scheme can reach up to 9 dB for BER $= 10^{-4}$ and a very large scale system setup. Finally, in order to validate the obtained simulation results, we analyze the Average Bit Error Probability (ABEP) using the well known union bound method [34]. We emphasize that the derivation of the ABEP is conducted both for the CD and DD algorithms at the RNs.

The rest of the paper is organized as follows: Section II presents the system model of the proposed architecture when the RNs employ CD. Section III concentrates at the SN and proposes two precoding methods that take into account I-CSIT at the SN. In addition, the same section provides a mathematical analysis of the instantaneous and average transmit power of the SN. In order to avoid coordination between the RNs, in Section IV we propose a DD algorithm which can be independently executed by the RNs. The derivation of the ABEP of our architecture is presented in Section V. Section VI provides results that evaluate the performance of the proposed scheme and precoding methods. Finally, we conclude this paper in Section VII.

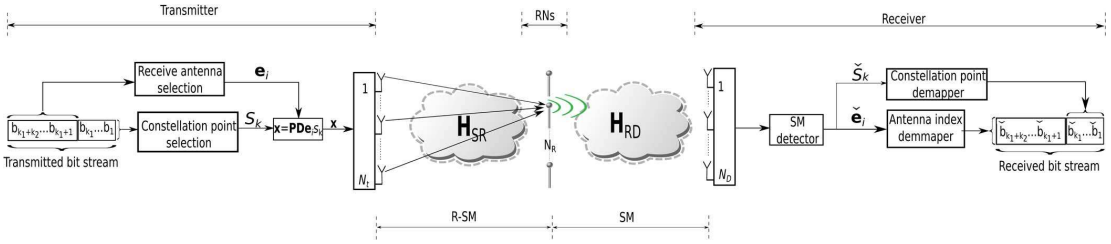


Fig. 1. System model of DH-HSM.

Notation: In the following, lowercase bold letters denote vectors and uppercase bolt letters denote matrices. $(\cdot)^T$, $(\cdot)^H$, $\mathrm{tr}(\cdot)$ and $\mathbf{A}^{1/2}$ denote transpose, Hermitian transpose, matrix trace and the square root of $\mathbf{A}$, respectively. The Kronecker product is denoted as $\otimes$. $\|\cdot\|_2$ represents the Euclidean norm, while $\|\cdot\|_F$ stands for the Frobenius norm. $\mathbf{diag}(a_1,\ldots,a_n)$ represents a diagonal matrix whose main diagonal includes the elements $a_1,\ldots,a_n$. $\mathrm{E}[\cdot]$ is the mean value of a Random Variable (RV). A complex Gaussian distribution with mean m and variance σ_C^2 is represented as $\mathcal{CN}(m,\sigma_C^2)$, where its real and imaginary part are independent and identically distributed (i.i.d.) Gaussian RV with distribution $\mathcal{N}(m,(\sigma_C^2/2))$. $\mathrm{Re}\{\cdot\}$ denotes the real part of a complex number or matrix.

II. System Model

Fig. 1 shows the system model of the proposed DH-HSM architecture. We consider a Base Station (BS) equipped with N_t antennas which acts as the SN. Furthermore, we assume that our architecture includes N_R single antenna RNs connected with an error-free backhaul link. *Later in Section IV, we extend our architecture in a distributed framework where there is no backhaul connection between the RNs*. The DN can be considered as a single or multi-antenna node (with N_D antennas).

In addition, we assume that there is no direct link between the SN and the DN due to the poor channel condition. The SN acquires CSIT using either the reciprocity principle when it is applicable or using a low-rate feedback link from the relays. The DN acquires CSI using a training sequence transmitted from the RNs. Especially, when Time Division Duplexing (TDD) is employed, DH-HSM is able to supply the SN and the DN with Channel State Information (CSI) by transmitting a single training sequence from the RNs. In this section, we assume Perfect-CSIT (P-CSIT) at DN for the purpose of system presentation and in Section III we modify our architecture in order to operate under different practical scenarios of I-CSIT at the SN.

We configure our system with $N_t \geq N_R$ and N_R equal to a power of 2. Under the assumptions that the transmitter employs a linear precoder and the wireless channel is flat quasi-static, the system equation of the first hop is expressed in a matrix form as:

$$\mathbf{y}_{\mathrm{R}} = \mathbf{H}_{\mathrm{SR}}\mathbf{P}\mathbf{D}\mathbf{x} + \mathbf{w}_{\mathrm{SR}}. \tag{1}$$

In (1), $\mathbf{y}_{\mathrm{R}}$ represents the $N_R \times 1$ received signal at the RNs in vector form. $\mathbf{H}_{\mathrm{SR}} \in \mathcal{C}^{N_R \times N_t}$ represents the channel matrix which is assumed to distribute as $\mathbf{H}_{\mathrm{SR}} \sim \mathcal{CN}(\mathbf{0},\mathbf{I})$. The $N_t \times N_R$ linear precoder is denoted as $\mathbf{P}$. $\mathbf{D}$ is a $N_R \times N_R$ diagonal matrix which is used to ensure that the transmitted power is not amplified by the precoder. The i-th element of the main diagonal of $\mathbf{D}$ is equal to $d_i = \sqrt{1/\|\mathbf{p}_i\|_2^2}$, where $\mathbf{p}_i$ is the i-th column of $\mathbf{P}$. In this way, each column of the normalized precoder $\mathbf{P}_{\mathrm{n}} = \mathbf{P}\mathbf{D}$ is constrained to have unity power. Furthermore, $\mathbf{x}$ represents the transmitting symbol vector. In this paper, we set the elements of $\mathbf{x}$ such that $\mathrm{E}_{\mathbf{x}}[\|\mathbf{x}\|_2^2] = 1$. Finally, $\mathbf{w}_{SR} \in \mathcal{C}^{N_D}$ denotes the i.i.d. additive complex Gaussian noise with $\mathbf{w}_{SR_i} \sim \mathcal{CN}(0,\sigma_{w_{SR}}^2)$. We note that due to the use of the linear precoder $\mathbf{P}$, the DN requires multiple (N_t) RF chains.

In this paper, under the assumption of P-CSIT at the SN, we focus on two types of linear precoding: i) ZF and ii) MMSE. In traditional MIMO schemes, ZF (Bezout) precoding is a suboptimal method that offers a good trade-off between complexity and performance [35]. The main characteristic of ZF is the total elimination of the Inter-Channel Interference (ICI) at the receiver $(\mathbf{H}_{\mathrm{SR}}\mathbf{P}_{\mathrm{ZF}} = \mathbf{I}_{N_R,N_R})$. The ZF precoding matrix is just the pseudo-inverse of the channel matrix $\mathbf{H}_{\mathrm{SR}}$. Using the Singular Value Decomposition (SVD), the precoding matrix is equal to

$$\mathbf{P}_{\mathrm{ZF}} = \mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^H, \tag{2}$$

where $\mathbf{H}_{\mathrm{SR}} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$. Both $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices with size $N_R \times N_R$ and $N_t \times N_t$, respectively. $\mathbf{\Sigma}$ is a $N_R \times N_t$ diagonal matrix whose main diagonal represents the r real singular values σ_i of $\mathbf{H}_{\mathrm{SR}}$. Here, r denotes the rank of the channel matrix $\mathbf{H}_{\mathrm{SR}}$. A different way to express the ZF precoder is given below:

$$\mathbf{P}_{\mathrm{ZF}} = \mathbf{H}_{\mathrm{SR}}^H \left(\mathbf{H}_{\mathrm{SR}}\mathbf{H}_{\mathrm{SR}}^H\right)^{-1}. \tag{3}$$

An alternative method, suitable for rank-deficient channel matrices, is MMSE (Regularized ZF) precoding. The MMSE is given as:

$$\mathbf{P}_{\mathrm{MMSE}} = \mathbf{H}_{\mathrm{SR}}^H \left[\mathbf{H}_{\mathrm{SR}}\mathbf{H}_{\mathrm{SR}}^H + \mu\mathbf{I}_{N_R,N_R}\right]^{-1}, \tag{4}$$

where μ is the regularization factor. Usually, μ is defined as $\mu = N_t\sigma_w^2/P_{\mathbf{s}} = N_t/\mathrm{SNR}$ [36]. In traditional MIMO, MMSE precoding outperforms ZF precoding in low SNRs and approaches the ZF performance in high SNRs.

In order to form a R-SM transmission mechanism from the SN to the RNs, such as [18], [19], we set the transmitted signal vector $\mathbf{x} \widehat{=} \mathbf{x}(i,s_k) = \mathbf{e}_i s_k$. Where $\mathbf{e}_i = [0,\ldots,0,1,0,\ldots,0]^T$,

$i = 1, \ldots, N_R$, is a $N_R \times 1$ vector, with all of its elements equal to zero except the i-th element which is equal to 1. The value of i corresponds to the index of the receiving RN. The indices i are allocated to the RNs in the Medium Access Layer (MAC). $s_k \in \{s_1, \ldots, s_M\}$ denotes the transmitted symbol selected from a conventional constellation. Hence, depending on the bits to be transmitted, $k_1 = \log_2 N_R$ bits are used for the selection of i and $k_2 = \log_2 M$ bits are encoded using the constellation point s_k. In this way, $k_1 + k_2$ bits are transmitted from the SN to the RNs in a symbol period (first hop).

Under the assumption of ZF precoding using P-CSIT at the SN, the observation at each of the RNs is given as:

$$\mathbf{y}_R = \mathbf{D}\mathbf{x} + \mathbf{w}_{SR}. \tag{5}$$

If we take into account the sparsity of $\mathbf{x}$, (5) is written as:

$$\begin{aligned} y_{R_j} &= d_j s_k + w_{SR_j}, \quad j = i, \\ y_{R_j} &= w_{SR_j}, \qquad\quad j \neq i, \end{aligned} \tag{6}$$

where $j = 1, \ldots, N_R$.

Due to the fact that the RNs communicate via a backhaul-link, we are able to employ the following centralized low-complexity ML detector:

$$(\hat{i}, \hat{s}_k) = \arg\min_{i, s_k} d_i^2 |s_k|^2 - 2 d_i \mathrm{Re}\left\{ y_{R_i^*} s_k \right\}. \tag{7}$$

As we can see from (7), the CSI that the R-SM detector requires is d_i, $i = 1, \ldots, N_R$. We note that (7) ensures the activation of a single RN during the second hop.

During the second slot, we apply a SM-like transmission at which only the receiving RN from the previous slot is activated. This RN acts as a single antenna in a conventional co-located SM system. In this way, the system equation of the RNs–DN link is given as:

$$\mathbf{y}_D = \mathbf{H}_{RD}\mathbf{x}(\hat{i}, \hat{s}_k) + \mathbf{w}_{RD}, \tag{8}$$

where $\mathbf{H}_{\mathrm{RD}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ stands for the $N_D \times N_R$ RNs–DN wireless channel. $\mathbf{x}(\hat{i}, \hat{s}_k) = \mathbf{e}_{\hat{i}} \hat{s}_k$ is the $N_R \times 1$ vectorized detected symbol from the previous slot. Moreover, $\mathbf{w}_{RD} \in \mathcal{C}^{N_D}$ is the i.i.d. additive complex Gaussian noise with $\mathbf{w}_{RD_i} \sim \mathcal{CN}(0, \sigma^2_{w_{RD}})$. Finally, $\mathbf{y}_D$ stands for the received signal at the DN. The optimal ML detection for SM is given in [11]

$$(\check{i}, \check{s}_k) = \arg\min_{i, s_k} \|\mathbf{y}_D - \mathbf{H}_{RD}\mathbf{e}_i s_k\|_2^2. \tag{9}$$

Hence, in this way, we form a dual-hop architecture which conveys $k_t = (k_1 + k_2)/2$ bits per symbol period (bpsp). We note that $k_1 + k_2$ is divided by two, because the transmitted bitstream requires two symbol periods to reach the DN.

A special case of the above transmission mechanism, called Dual-Hop Hybrid Space Shift Keying (DH-HSSK), is formed if we set $s_k = 1$. In such a scenario, the information is conveyed using only the index i of the receiving RN.

III. Linear Precoding at the SN Using Practical Scenarios of I-CSIT

When P-CSIT is available at the SN, the design of the precoding matrix $\mathbf{P}$ is trivial, as in Section II. Our aim in the next two subsection is to design the precoding matrix of (1) using I-CSIT.

A. Imperfect Channel Knowledge at the Transmitter

In practice, CSI is acquired by the receiver using a pilot sequence. Under the assumption that the channel does not change during one block of symbols, Channel State Information at the Receiver (CSIR) may be considered accurate. When the channel is rapidly varying, even CSIR is subjected to imperfections. At the transmitter side, CSI is acquired either using a low rate feedback from the receiver or using the reciprocity principle where is applicable (pure TDD). In the general case, channel imperfections at the transmitter side may occur because of: i) the channel estimation errors; ii) the quantization errors when a reverse feedback link is employed; and iii) the outdated version of CSIT (time varying channel).

Clearly, in practical scenarios the acquisition of perfect (accurate) CSI is either a very expensive or even an unrealistic process. Hence, there is a high need for designs that take into account channel imperfections. A widely used channel model, that captures channel imperfections, expresses the channel matrix as [37],

$$\mathbf{H}_{\mathrm{SR}} = \bar{\mathbf{H}}_{\mathrm{SR}} + \widetilde{\mathbf{H}}_{\mathrm{SR}}. \tag{10}$$

In (10), $\bar{\mathbf{H}}_{\mathrm{SR}}$ represents the long term channel evolution which can be accurately acquired by the transmitter. Moreover, $\widetilde{\mathbf{H}}_{\mathrm{SR}}$ denotes the channel for which only some kind of statistical or worst-case knowledge is considered possible. A common assumption is to model the long term channel evolution as $\bar{\mathbf{H}}_{\mathrm{SR}} \sim \mathcal{CN}(\mathbf{0}, \sigma^2_{\bar{\mathbf{H}}_{\mathrm{SR}}}\mathbf{I})$, while $\widetilde{\mathbf{H}}_{\mathrm{SR}} \sim \mathcal{CN}(\mathbf{0}, \sigma^2_{\widetilde{\mathbf{H}}_{\mathrm{SR}}}\mathbf{I})$. We note that each element of $\bar{\mathbf{H}}_{\mathrm{SR}}$ remains constant over a block of symbols. However, the elements of $\widetilde{\mathbf{H}}_{\mathrm{SR}}$ change every symbol period. $\mathbf{H}_{\mathrm{SR}}$ denotes the instantaneous channel matrix. Finally, in order to avoid any power amplification due to the wireless channel, we assume $\sigma^2_{\bar{\mathbf{H}}_{\mathrm{SR}}} + \sigma^2_{\widetilde{\mathbf{H}}_{\mathrm{SR}}} = 1$.

In the following two subsections, we design the precoding matrix $\mathbf{P}$ at the SN using two methods from optimization theory, Stochastic Robust Approximation (SRA) and Worst-Case Robust Approximation (WCRA) [38]. SRA follows a statistical approach and solves the resulting minimization problems over their expectation. In contrast, WCRA is a worst-case method and solves the same minimization problems assuming that $\|\widetilde{\mathbf{H}}_{\mathrm{SR}}\|_2 \leq \alpha$. We note that both approaches employ I-CSIT in a statistical or scalar form which does not change rapidly over time.

B. Objective Function

In this subsection, we present the objective function that we minimize in order to design the precoding matrix $\mathbf{P}$ using I-CSIT. When P-CSIT is available, the ZF precoding matrix of (1) is just the pseudo-inverse matrix of $\mathbf{H}_{\mathrm{SR}}$. An alternative

way to obtain the ZF precoder is to solve the following N_R minimization problems:

$$\min_{\mathbf{p}_l} \|\mathbf{H}_{\mathrm{SR}}\mathbf{p}_l - \mathbf{e}_l\|_2^2, \quad \forall\, l = 1, \ldots, N_R, \tag{11}$$

where $\mathbf{e}_l$ is the l-th column of the identity matrix $\mathbf{I}_{N_R,N_R} = [\mathbf{e}_1, \ldots, \mathbf{e}_{N_R}]$.

C. ZF-Like Precoding Based on SRA

When SRA is employed, the minimization problems of (11) are solved over their expectation [38]

$$\min_{\mathbf{p}_l} \mathrm{E}\left[\|\mathbf{H}_{\mathrm{SR}}\mathbf{p}_l - \mathbf{e}_l\|_2^2\right], \quad \forall\, l = 1, \ldots, N_R. \tag{12}$$

If we use the fact that the MIMO channel can be written as in (10), where $\mathrm{E}[\widetilde{\mathbf{H}}_{\mathrm{SR}}] = \mathbf{0}$, (12) can be reformulated as:

$$\min_{\mathbf{p}_l} \left\{\|\bar{\mathbf{H}}_{\mathrm{SR}}\mathbf{p}_l - \mathbf{e}_l\|_2^2 + \mathbf{p}_l^H \mathrm{E}\left[\widetilde{\mathbf{H}}_{\mathrm{SR}}^H \widetilde{\mathbf{H}}_{\mathrm{SR}}\right]\mathbf{p}_l\right\}, \tag{13}$$

which is obviously a convex optimization problem (sum of quadratic functions). Thus, the solution can be achieved by setting its gradient equal to zero. In this case, the analytical solution is written as:

$$\mathbf{p}_l = \left[\bar{\mathbf{H}}_{\mathrm{SR}}^H \bar{\mathbf{H}}_{\mathrm{SR}} + \mathrm{E}\left[\widetilde{\mathbf{H}}_{\mathrm{SR}}^H \widetilde{\mathbf{H}}_{\mathrm{SR}}\right]\right]^{-1} \bar{\mathbf{H}}_{\mathrm{SR}}^H \mathbf{e}_l, \tag{14}$$

or

$$\mathbf{P}_{\mathrm{SRA}} = \left[\bar{\mathbf{H}}_{\mathrm{SR}}^H \bar{\mathbf{H}}_{\mathrm{SR}} + \mathrm{E}\left[\widetilde{\mathbf{H}}_{\mathrm{SR}}^H \widetilde{\mathbf{H}}_{\mathrm{SR}}\right]\right]^{-1} \bar{\mathbf{H}}_{\mathrm{SR}}^H, \tag{15}$$

in a collective matrix form. A careful look in $\mathrm{E}[\widetilde{\mathbf{H}}_{\mathrm{SR}}^H \widetilde{\mathbf{H}}_{\mathrm{SR}}]$ reveals its structured form. For example, when $\widetilde{\mathbf{H}}_{\mathrm{SR}}$ is uncorrelated, we have:

$$\mathrm{E}\left[\widetilde{\mathbf{H}}_{\mathrm{SR}}^H \widetilde{\mathbf{H}}_{\mathrm{SR}}\right] = N_R \sigma^2_{\widetilde{\mathbf{H}}_{\mathrm{SR}}} \mathbf{I}_{N_t,N_t}. \tag{16}$$

Thus, the computation of the previous quantity requires only the knowledge of $\sigma^2_{\widetilde{\mathbf{H}}_{\mathrm{SR}}}$, which is scalar.

D. ZF-Like Precoding Based on WCRA

When WCRA is utilized, the MIMO wireless channel is expressed as the non-empty and bounded set $\Phi \subseteq \mathcal{C}^{N_R,N_t}$. Here, Φ represents all the possible values of the channel matrix $\mathbf{H}_{\mathrm{SR}}$. Given a feasible precoding vector $\mathbf{p}_l$, the worst case error can be formulated as $e_{\mathrm{wc}}(\mathbf{p}_l) = \sup[\|\mathbf{H}_{\mathrm{SR}}\mathbf{p}_l - \mathbf{e}_l\|_2 | \mathbf{H} \in \Phi]$. Our aim in this subsection is to design a precoder that minimizes the worst case error $e_{\mathrm{wc}}(\mathbf{p}_l)$. Under this aim, the minimization problem can be formulated as:

$$\min_{\mathbf{p}_l} \sup\left[\|\mathbf{H}_{\mathrm{SR}}\mathbf{p}_l - \mathbf{e}_l\|_2 | \mathbf{H}_{\mathrm{SR}} \in \Phi\right] \quad \forall\, l = 1, \ldots, N_R. \tag{17}$$

In this paper, we employ the Norm Bound Error (NBE) method from [38]. In NBE, the uncertainty of $\widetilde{\mathbf{H}}_{\mathrm{SR}}$ is considered within a norm ball of radius α. Thus, the set Φ is written as:

$$\Phi = \left\{\mathbf{H}_{\mathrm{SR}} = \bar{\mathbf{H}}_{\mathrm{SR}} + \widetilde{\mathbf{H}}_{\mathrm{SR}} | \|\widetilde{\mathbf{H}}_{\mathrm{SR}}\|_2 \leq \alpha\right\},$$

where $\alpha > 0$. Let

$$e_{\mathrm{wc}}^{\mathrm{NBE}}(\mathbf{p}_l) = \sup\left\{\|\bar{\mathbf{H}}_{\mathrm{SR}}\mathbf{p}_l - \mathbf{e}_l + \widetilde{\mathbf{H}}_{\mathrm{SR}}\mathbf{p}_l\|_2 | \, \|\widetilde{\mathbf{H}}_{\mathrm{SR}}\|_2 \leq \alpha\right\}$$

be the worst-case error given the precoding vector $\mathbf{p}_l$. After some arithmetic manipulations, it is shown that $e_{\mathrm{wc}}^{\mathrm{NBE}}(\mathbf{p}_l)$ is equal to

$$e_{\mathrm{wc}}^{\mathrm{NBE}}(\mathbf{p}_l) = \|\bar{\mathbf{H}}_{\mathrm{SR}}\mathbf{p}_l - \mathbf{e}_l\|_2 + \alpha\|\mathbf{p}_l\|_2$$

and it is attained for $\widetilde{\mathbf{H}}_{\mathrm{SR}} = \alpha\mathbf{u}\mathbf{v}^H$, where

$$\mathbf{u} = \frac{\bar{\mathbf{H}}_{\mathrm{SR}}\mathbf{p}_l - \mathbf{e}_l}{\|\bar{\mathbf{H}}_{\mathrm{SR}}\mathbf{p}_l - \mathbf{e}_l\|_2} \text{ and } \mathbf{v} = \frac{\mathbf{p}_l}{\|\mathbf{p}_l\|_2},$$

given that

$$\bar{\mathbf{H}}_{\mathrm{SR}}\mathbf{p}_l - \mathbf{e}_l \neq \mathbf{0} \text{ and } \mathbf{p}_l \neq \mathbf{0}.$$

Thus, the minimization problem of (17) can be reformulated as:

$$\min_{\mathbf{p}_l} \|\bar{\mathbf{H}}_{\mathrm{SR}}\mathbf{p}_l - \mathbf{e}_l\|_2 + \alpha\|\mathbf{p}_l\|_2. \tag{18}$$

The minimization problems of (18) can be transformed to:

$$\begin{aligned} \min_{\{t_1,t_2\}} \quad & t_1 + at_2 \\ \text{subject to} \quad & \|\bar{\mathbf{H}}_{\mathrm{SR}}\mathbf{p}_l - \mathbf{e}_l\|_2 \leq t_1, \ \|\mathbf{p}_l\|_2 \leq t_2, \end{aligned} \tag{19}$$

which is solved as a Second Order Cone Programming (SOCP) problem using the interior point method [38]. Alternatively, (18) can be transformed in a TR form as:

$$\min_{\mathbf{p}_l} \|\bar{\mathbf{H}}_{\mathrm{SR}}\mathbf{p}_l - \mathbf{e}_l\|_2^2 + \beta\|\mathbf{p}_l\|_2^2, \tag{20}$$

for some value of β [38].

Again, (20) is convex because $e_{\mathrm{wc}}^{\mathrm{NBE}}(\mathbf{p}_l)$ is the sum of quadratic functions. Thus, the solution can be reached using the gradient condition and is written as:

$$\mathbf{p}_l = \left[\bar{\mathbf{H}}_{\mathrm{SR}}^H \bar{\mathbf{H}}_{\mathrm{SR}} + \beta\mathbf{I}\right]^{-1} \bar{\mathbf{H}}_{\mathrm{SR}}^H \mathbf{e}_l. \tag{21}$$

The collective matrix form of the WCRA precoder is expressed as:

$$\mathbf{P}_{\mathrm{WCRA}} = \left[\bar{\mathbf{H}}_{\mathrm{SR}}^H \bar{\mathbf{H}}_{\mathrm{SR}} + \beta\mathbf{I}\right]^{-1} \bar{\mathbf{H}}_{\mathrm{SR}}^H. \tag{22}$$

E. Precoding in the Presence of Transmit and Receive Space Correlations

The problems of (13) and (20) are different forms of Tikhonov Regularization (TR) [38]. A valuable property of TR theory is that it does not pose any rank restriction on the involved matrices $\bar{\mathbf{H}}_{\mathrm{SR}}$ and $\widetilde{\mathbf{H}}_{\mathrm{SR}}$ as long as the matrices $\bar{\mathbf{H}}_{\mathrm{SR}}^H \bar{\mathbf{H}}_{\mathrm{SR}} + \mathrm{E}[\widetilde{\mathbf{H}}_{\mathrm{SR}}^H \widetilde{\mathbf{H}}_{\mathrm{SR}}]$ and $\bar{\mathbf{H}}_{\mathrm{SR}}^H \bar{\mathbf{H}}_{\mathrm{SR}} + \beta\mathbf{I}$ are positive

definite [38]. Thus, our analytical solutions of (15) and (22) may enjoy the additional merit of being applicable to spatially correlated channels.

In this paper, we employ the Kronecker correlation model [39]. According to this correlation model, the MIMO channel can be rewritten as:

$$\mathbf{H}_{\mathrm{SR}} = \mathbf{R}_R^{1/2}\mathbf{H}_w\left(\mathbf{R}_T^{1/2}\right)^T, \tag{23}$$

where $\mathbf{H}_w \sim \mathcal{CN}(\mathbf{0},\mathbf{I})$. $\mathbf{R}_T$ and $\mathbf{R}_R$ represent the transmit and receive spatial correlation matrices, respectively. In this case the wireless channel is distributed as $\mathbf{H}_{\mathrm{SR}} \sim \mathcal{CN}(\mathbf{0},\mathbf{R}_T \otimes \mathbf{R}_R)$. Usually, the entries of the spatial correlation matrices $\mathbf{R}_R$ and $\mathbf{R}_T$ are generated using an exponential model with $R_T(i,j) = \rho_t^{|i-j|}$ and $R_R(i,j) = \rho_r^{|i-j|}$, where $0 \leq \rho_t,\ \rho_r \leq 1$. Values of ρ_t and ρ_r close to 0 mean low correlation, whereas values close to 1 mean high correlations.

If we combine the Kronecker correlation model of (23) with the model of I-CSIT described in Section III-A, the MIMO channel is expressed as:

$$\mathbf{H}_{\mathrm{SR}} = \mathbf{R}_R^{1/2}\left(\bar{\mathbf{H}}_{\mathrm{SR}} + \widetilde{\mathbf{H}}_{\mathrm{SR}}\right)\left(\mathbf{R}_T^{1/2}\right)^T. \tag{24}$$

Here, the matrix $\mathbf{R}_R^{1/2}\bar{\mathbf{H}}_{\mathrm{SR}}\left(\mathbf{R}_T^{1/2}\right)^T$ represents the fully known part of the channel and the matrix $\mathbf{R}_R^{1/2}\widetilde{\mathbf{H}}_{\mathrm{SR}}\left(\mathbf{R}_T^{1/2}\right)^T$ represents the channel uncertainty.

The inspection of (15) and (22) reveals that only the analytical form of the design based on SRA is affected by the correlated channel. This is because SRA requires the computation of $\mathrm{E}[\widetilde{\mathbf{H}}_{\mathrm{SR}}^H\widetilde{\mathbf{H}}_{\mathrm{SR}}]$. After some arithmetic manipulations this quantity is expressed as:

$$\mathrm{E}\left[\widetilde{\mathbf{H}}_{\mathrm{SR}}^H\widetilde{\mathbf{H}}_{\mathrm{SR}}\right] = \sigma^2_{\widetilde{\mathbf{H}}_{\mathrm{SR}}}\,\mathrm{tr}(\mathbf{R}_R)\mathbf{R}_T^{1/2}\left(\mathbf{R}_T^{1/2}\right)^T. \tag{25}$$

F. Instantaneous Transmit Power Analysis

In this subsection, we focus on the SN and study its transmission power. Given that $\mathrm{E}_{\mathbf{x}}[\|\mathbf{x}\|_2^2] = 1$, the instantaneous transmission power $\|\mathbf{PDx}\|_2^2$ is fully characterized by the quantity $P_{\mathrm{s}} = \|\mathbf{PD}\|_2^2$.

In (1), we employ the normalization matrix $\mathbf{D}$ in order to make sure that the transmission power at the SN is not amplified by the precoder $\mathbf{P}$ in adverse channel conditions. As shown in the following, the proposed precoders have a structure that imposes an indirect transmit power constraint. The practical value of this constraint is that we are able to set $\mathbf{D} = \mathbf{I}_{N_R,N_R}$, without amplifying the power transmission to an unacceptable level. Additionally, we can assume that $\mathbf{H}_{\mathrm{SR}}\mathbf{P} \approx \mathbf{I}_{N_R,N_R}$ (when $\mathbf{P}$ is the MMSE, or SRA, or WCRA precoder) in order to keep the complexity of the RNs low. In this way, there is no need for d_i to be known at the RNs. Hence, the detection algorithm of (7) is further simplified by setting $d_i = 1$. *Though, we emphasize that in the BER simulation results of Section VI the precoding designs of SRA and WCRA are normalized as described in Section II. We follow this choice for the sake of fair comparison with the other precoding methods (same transmission power).*

SRA and WCRA are forms of TR when I-CSIT is available. As a consequence, we expect both designs to offer a reduced transmit power [38]. Let us begin our proof by taking the SVD of $\bar{\mathbf{H}}_{\mathrm{SR}}$:

$$\bar{\mathbf{H}}_{\mathrm{SR}} = \bar{\mathbf{U}}\bar{\mathbf{\Sigma}}\bar{\mathbf{V}}^H. \tag{26}$$

Here, $\bar{\mathbf{U}}$ and $\bar{\mathbf{V}}$ are unitary matrices and $\bar{\mathbf{\Sigma}}$ is a diagonal matrix containing at its main diagonal the N_R singular values $\bar{\sigma}_i$ of $\bar{\mathbf{H}}_{\mathrm{SR}}$ (without loss of generality we assume that $\bar{\mathbf{H}}_{\mathrm{SR}}$ has a rank of N_R). If we plug (26) into (15) and after some arithmetic manipulations, the SRA precoder is written as:

$$\mathbf{P}_{\mathrm{SRA}} = \bar{\mathbf{V}}\left[\bar{\mathbf{D}} + N_R\sigma^2_{\widetilde{\mathbf{H}}_{\mathrm{SR}}}\mathbf{I}_{N_t,N_t}\right]^{-1}\bar{\mathbf{\Sigma}}^T\bar{\mathbf{U}}^H, \tag{27}$$

where $\bar{\mathbf{D}}$ is a $N_t \times N_t$ diagonal matrix equal to $\bar{\mathbf{D}} = \mathbf{diag}(\bar{\sigma}_1^2,\ldots,\bar{\sigma}_{N_R}^2)$. $\bar{\sigma}_i^2$, $i = 1,\ldots,N_R$ are the eigenvalues of $\bar{\mathbf{H}}_{\mathrm{SR}}$.

Taking into account that both $\bar{\mathbf{D}} + N_R\sigma^2_{\widetilde{\mathbf{H}}_{\mathrm{SR}}}\mathbf{I}_{N_t,N_t}$ and $\bar{\mathbf{\Sigma}}$ are diagonal matrices, we can further proceed with (27) and show that

$$\mathbf{P}_{\mathrm{SRA}} = \bar{\mathbf{V}}\bar{\mathbf{\Sigma}}_{\mathrm{SRA}}\bar{\mathbf{U}}^H, \tag{28}$$

where

$$\mathbf{\Sigma}_{\mathrm{SRA}} = \mathbf{diag}\left(\frac{\bar{\sigma}_1}{\bar{\sigma}_1^2 + N_R\sigma^2_{\widetilde{\mathbf{H}}_{\mathrm{SR}}}},\ldots,\frac{\bar{\sigma}_{N_R}}{\bar{\sigma}_{N_R}^2 + N_R\sigma^2_{\widetilde{\mathbf{H}}_{\mathrm{SR}}}}\right). \tag{29}$$

In order to evaluate the transmit power of the SRA precoder, we have to compute the Frobenius norm of the precoding matrix of (28)

$$\|\mathbf{P}_{\mathrm{SRA}}\|_F^2 = \mathrm{tr}\left(\mathbf{P}_{\mathrm{SRA}}^H\mathbf{P}_{\mathrm{SRA}}\right) = \mathrm{tr}\left(\bar{\mathbf{U}}\mathbf{\Sigma}_{\mathrm{SRA}}^T\mathbf{\Sigma}_{\mathrm{SRA}}\bar{\mathbf{U}}^H\right). \tag{30}$$

Using the argument that the trace operator is invariant under cyclic permutations $\mathrm{tr}(\mathbf{ABC}) = \mathrm{tr}(\mathbf{CAB})$, we can show that

$$\|\mathbf{P}_{\mathrm{SRA}}\|_F^2 = \sum_{i=1}^{N_R}\left(\frac{\bar{\sigma}_i}{\bar{\sigma}_i^2 + N_R\sigma^2_{\widetilde{\mathbf{H}}_{\mathrm{SR}}}}\right)^2. \tag{31}$$

In order to quantify the maximum of $\|\mathbf{P}_{\mathrm{SRA}}\|_F^2$, we focus on each component

$$\theta_i = \frac{\bar{\sigma}_i}{\bar{\sigma}_i^2 + N_R\sigma^2_{\widetilde{\mathbf{H}}_{\mathrm{SR}}}} \tag{32}$$

of the summation of (31). Given that $\bar{\sigma}_i$ and $\sigma^2_{\widetilde{\mathbf{H}}_{\mathrm{SR}}}$ take positive values, it is clear to see that θ_i: i) is close to zero when $\bar{\sigma}_i \to 0$; ii) it attains the maximum value of $1/(1 + N_R\sigma^2_{\widetilde{\mathbf{H}}_{\mathrm{SR}}}) < 1$ when $\bar{\sigma}_i = N_R\sigma^2_{\widetilde{\mathbf{H}}_{\mathrm{SR}}}$; and iii) approaches zero when $\bar{\sigma}_i \to +\infty$. Hence, the instantaneous transmit power of $\mathbf{P}_{\mathrm{SRA}}$ is always constrained by $\|\mathbf{P}_{\mathrm{SRA}}\|_F^2 < N_R$, since $\bar{\sigma}_i/(\bar{\sigma}_i^2 + N_R\sigma^2_{\widetilde{\mathbf{H}}_{\mathrm{SR}}}) < 1$ for every positive value of $\bar{\sigma}_i$ and $\sigma^2_{\widetilde{\mathbf{H}}_{\mathrm{SR}}}$.

Again, following the same steps, we can express the WCRA precoder as:

$$\mathbf{P}_{\mathrm{WCRA}} = \bar{\mathbf{V}}\bar{\mathbf{\Sigma}}_{\mathrm{WCRA}}\bar{\mathbf{U}}^H, \tag{33}$$

where

$$\mathbf{\Sigma}_{\mathrm{WCRA}} = \mathbf{diag}\left(\frac{\bar{\sigma}_1}{\bar{\sigma}_1^2+\beta}, \ldots, \frac{\bar{\sigma}_{N_R}}{\bar{\sigma}_{N_R}^2+\beta}\right). \quad (34)$$

The transmit power of WCRA is expressed as:

$$\|\mathbf{P}_{\mathrm{WCRA}}\|_F^2 = \sum_{i=1}^{N_R}\left(\frac{\bar{\sigma}_i}{\bar{\sigma}_i^2+\beta}\right)^2 < N_R. \quad (35)$$

Clearly, $\|\mathbf{P}_{\mathrm{WCRA}}\|_F^2$ is also constrained because $(\bar{\sigma}_i/\bar{\sigma}_i^2 + \beta) < 1$ for every $\{\bar{\sigma}_i, \beta\} > 0$.

A careful inspection of (31) and (35) reveals that the channel uncertainty (in the form of $\sigma^2_{\tilde{\mathbf{H}}_{\mathrm{SR}}}$ for SRA and in the form of β for WCRA) acts as a regularization parameter which adjusts the transmitted power. When the channel uncertainty approaches to zero, both SRA and WCRA reduce to ZF precoding and the transmitted power is maximized up to the transmitted power of ZF precoding. In contrast, when the channel uncertainty is increased, the level of power transmission imposed by the transmitter is decreased. This means that when the channel knowledge at the transmitter becomes more inaccurate, the transmitter reduces the transmitted power in order to avoid further degradation.

In the next few lines, we express without proof (due to space limitation) the transmit power behavior of ZF and MMSE precoding when I-CSIT is available at the SN. Using a similar approach as previously, it is shown that the transmit power of ZF precoding is:

$$\|\mathbf{P}_{\mathrm{ZF}}\|_F^2 = \sum_{i=1}^{N_R}\frac{1}{\bar{\sigma}_i} \quad (36)$$

and the transmit power of MMSE precoding is:

$$\|\mathbf{P}_{\mathrm{MMSE}}\|_F^2 = \sum_{i=1}^{N_R}\left(\frac{\bar{\sigma}_i}{\bar{\sigma}_i{}^2+\mu}\right)^2. \quad (37)$$

From (36), it is clear that the power transmission of ZF precoding is increased to unacceptable high levels when the channel is in deep fade ($\bar{\sigma}_i^2 \to 0$). In contrast, for MMSE precoding it can be shown that the power transmission is constrained by $\|\mathbf{P}_{\mathrm{MMSE}}\|_F^2 < N_R$ because $\bar{\sigma}_i/(\bar{\sigma}_i^2+\mu) < 1$, $i = 1, \ldots, N_R$. A similar proof for the ZF and MMSE precoders is attained when the SN employs P-CSIT. In this case, the power transmission of the ZF and MMSE is given from (36) and (37) by replacing $\bar{\sigma}_i$ with σ_i, where the eigenvalues of $\mathbf{H}_{\mathrm{SR}}$ are denoted as σ_i^2.

Finally, the inspection of (31), (35), and (37) shows that there is a relation between the transmission power of different precoding methods. It is easy to demonstrate, using the structure of the power transmission of each precoder, that MMSE imposes higher instantaneous power transmission than: i) SRA when $\sigma^2_{\tilde{\mathbf{H}}_{\mathrm{SR}}} > \mu/N_R$, and ii) WCRA when $\beta > \mu$.

G. Average Transmit Power Analysis

In this subsection, we provide the analytical derivation of the average precoding power $\bar{P}_{\mathrm{s}}$ of the methods employed in this work when: i) I-CSIT is available and ii) $\mathbf{D} = \mathbf{I}_{N_R,N_R}$. The inspection of (31), (35)–(37) reveals that the instantaneous power of each precoder has the following structured form:

$$\|\mathbf{P}\|_F^2 = \sum_{i=1}^{N_R}\left(\frac{\bar{\sigma}_i}{\bar{\sigma}_i{}^2+\vartheta}\right)^2 = \sum_{i=1}^{N_R}\left(\frac{\sqrt{\bar{\lambda}_i}}{\bar{\lambda}_i+\vartheta}\right)^2. \quad (38)$$

where, depending on the precoding method, ϑ takes the following values: i) $\vartheta = 0$ for ZF; ii) $\vartheta = \mu$ for MMSE; iii) $\vartheta = N_R\sigma^2_{\tilde{\mathbf{H}}_{\mathrm{SR}}}$ for SRA; and iv) $\vartheta = \beta$ for WCRA. In (38), for notational convenience, we substitute $\bar{\sigma}_i{}^2$ with $\bar{\lambda}_i$.

The analytical derivation of the average power of all precoding methods is given as:

$$\bar{P}_s = \mathrm{E}\left[\|\mathbf{P}\|_F^2\right] = \mathrm{E}_{\bar{\boldsymbol{\tau}}}\left[\sum_{i=1}^{N_R}\left(\frac{\sqrt{\bar{\lambda}_i}}{\bar{\lambda}_i+\vartheta}\right)^2\right], \quad (39)$$

where $\bar{\boldsymbol{\tau}} = [\bar{\lambda}_1, \ldots, \bar{\lambda}_{N_R}]^T$. It is not difficult to see that the random vector $\bar{\boldsymbol{\tau}}$ contains the unordered eigenvalues of the Wishart matrix $\bar{\mathbf{W}} = \bar{\mathbf{H}}_{\mathrm{SR}}\bar{\mathbf{H}}_{\mathrm{SR}}^H \sim \mathcal{CW}(N_R, \mathbf{\Sigma}_{\bar{\mathbf{H}}_{\mathrm{SR}}})$, where $\mathcal{CW}(N_R, \mathbf{\Sigma}_{\bar{\mathbf{H}}_{\mathrm{SR}}})$ denotes a central complex Wishart distribution as defined in [40, eq. 2.6]. $\mathbf{\Sigma}_{\bar{\mathbf{H}}_{\mathrm{SR}}}$ is the covariance matrix of $\bar{\mathbf{H}}_{\mathrm{SR}}$. In the following, we derive the expectation of (39) when the SN–RNs channel is : i) uncorrelated and ii) transmit correlated ($\mathbf{R}_R = \mathbf{I}_{N_R,N_R}$). The latter case corresponds to the scenario where the transmit antennas belongs to a single transmitter, while the receive antennas belong to RNs placed sufficiently apart.

1) Uncorrelated SN–RNs Channel: When the SN–RNs is uncorrelated and distributed as described in Section III-A, the covariance matrix of $\bar{\mathbf{H}}_{\mathrm{SR}}$ equals to $\mathbf{\Sigma}_{\bar{\mathbf{H}}_{\mathrm{SR}}} = \sigma^2_{\bar{\mathbf{H}}_{\mathrm{SR}}}\mathbf{I}$. Given that $\bar{\mathbf{H}}_{\mathrm{SR}} = \sqrt{\sigma^2_{\bar{\mathbf{H}}_{\mathrm{SR}}}}\mathbf{H}_w$, where $\mathbf{H}_w \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$, the eigenvalues $\bar{\lambda}_i$ of $\bar{\mathbf{H}}_{\mathrm{SR}}$ can be rewritten as $\bar{\lambda}_i = \sqrt{\sigma^2_{\bar{\mathbf{H}}_{\mathrm{SR}}}}\lambda_i$. Here, λ_i, $i = 1, \ldots, N_R$, represent the eigenvalues of $\mathbf{H}_w$.

If we take into account the previous relation and set : i) $\bar{\boldsymbol{\tau}} = \sqrt{\sigma^2_{\bar{\mathbf{H}}_{\mathrm{SR}}}}\boldsymbol{\tau}$ and ii) $\phi = \vartheta/\sqrt{\sigma^2_{\bar{\mathbf{H}}_{\mathrm{SR}}}}$, after some manipulations, (39) is transformed to

$$\bar{P}_s = \frac{1}{\sigma^2_{\bar{\mathbf{H}}_{\mathrm{SR}}}}\mathrm{E}_{\boldsymbol{\tau}}\left[\sum_{i=1}^{N_R}\frac{\lambda_i}{(\lambda_i+\phi)^2}\right] = \frac{N_R}{\sigma^2_{\bar{\mathbf{H}}_{\mathrm{SR}}}}\mathrm{E}_{\lambda}\left[\frac{\lambda}{(\lambda+\phi)^2}\right]. \quad (40)$$

In (40), λ denotes the unordered eigenvalues of the complex central Wishart matrix $\mathbf{W} = \mathbf{H}_w\mathbf{H}_w^H$. The marginal probability density function (p.d.f.) of λ is given in [40, eq. 2.23] as:

$$f_{\lambda}(x) = \frac{1}{N_R}\left[\sum_{k=0}^{N_R-1}\frac{k!}{(k+\Delta)!}\left[L_k^{\Delta}(x)\right]^2 x^{\Delta}e^{-x}\right]H_0(x), \quad (41)$$

where $\Delta = N_t - N_R$, and

$$L_k^{\Delta}(x) = \sum_{i=0}^{k}(-1)^i\binom{k+\Delta}{k-i}\frac{x^i}{i!}. \quad (42)$$

Here, $L_k^{\Delta}(x)$ denotes the Laguerre polynomials of degree k. In addition, $H_0(x)$ is the Heaviside step function, for which it holds that $H_0(x)=0$ for $x<0$ and $H_0(x)=1$ for $x\geq 0$.

Thus, if we evaluate the expectation

$$\mathrm{E}_{\lambda}\left[\frac{\lambda}{(\lambda+\phi)^2}\right]=\int_0^{+\infty}\frac{x}{(x+\phi)^2}f_{\lambda}(x)dx \tag{43}$$

using (41) (in a lengthy procedure which is omitted due to space limitation), the average precoding power of (40) becomes:

$$\bar{P}_s=\frac{1}{\sigma^2_{\bar{\mathbf{H}}_{\mathrm{SR}}}}\sum_{k=0}^{N_R-1}\frac{k!}{(k+\Delta)!}\sum_{l=0}^{k}\sum_{m=0}^{k}\frac{(-1)^{l+m}}{l!m!}\binom{k+\Delta}{k-l}\times\binom{k+\Delta}{k-m}(n+1)!\phi^n e^{\phi}\left[\Gamma(-n,\phi)-\phi\Gamma(-n-1,\phi)\right]. \tag{44}$$

In (44), $n=l+m+\Delta$ and $\Gamma(\cdot,\cdot)$ denotes the incomplete gamma function defined in [41, p. 899].

2) Transmit Correlated SN–RNs Channel: The elaboration of (39) gives that:

$$\bar{P}_s=\mathrm{E}_{\bar{\boldsymbol{\tau}}}\left[\sum_{i=1}^{N_R}\left(\frac{\sqrt{\bar{\lambda}_i}}{\bar{\lambda}_i+\vartheta}\right)^2\right]=N_R\mathrm{E}_{\bar{\lambda}}\left[\frac{\bar{\lambda}}{(\bar{\lambda}+\vartheta)^2}\right], \tag{45}$$

where $\bar{\lambda}$ is the unordered eigenvalues of the complex central Wishart matrix $\bar{\mathbf{W}}=\bar{\mathbf{H}}_{\mathrm{SR}}\bar{\mathbf{H}}_{\mathrm{SR}}^H$. Clearly, the evaluation of (45) requires the marginal p.d.f. of $\bar{\lambda}$, which is derived below.

The presence of only transmit space correlation in the wireless channel determines that $\bar{\mathbf{H}}_{\mathrm{SR}}=\mathbf{H}_w(\bar{\mathbf{R}}_T^{1/2})^T$, where $\bar{\mathbf{R}}_T^{1/2}=\sqrt{\sigma^2_{\bar{\mathbf{H}}_{\mathrm{SR}}}}(\mathbf{R}_T^{1/2})^T$. Thus, $\bar{\mathbf{H}}_{\mathrm{SR}}$ is distributed according to the following complex Gaussian distribution $\mathbf{H}_{\mathrm{SR}}\sim\mathcal{CN}(\mathbf{0},\bar{\mathbf{R}}_T\otimes\mathbf{I}_{N_R,N_R})$. In this case, the complex central Wishart matrix $\bar{\mathbf{W}}$ is distributed as $\mathcal{CW}(N_R,\bar{\mathbf{R}}_T)$. For this reason, as it is shown in [42], the marginal p.d.f. of $\bar{\lambda}$ equals to:

$$f_{\bar{\lambda}}(x)=K\left[\sum_{i=1}^{N_R}\sum_{j=1}^{N_R}x^{j-1}\mathcal{D}(i,j)\left[\alpha_{\varrho+i}^{\varrho-1}e^{\frac{-x}{\alpha_{\varrho+i}}}-\sum_{l=1}^{\varrho}\sum_{k=1}^{\varrho}[\boldsymbol{\Psi}^{-1}]_{k,l}\alpha_{\varrho+i}^{k-1}\alpha_l^{\varrho-1}e^{\frac{-x}{\alpha_l}}\right]\right]H_0(x), \tag{46}$$

where $\varrho=N_t-N_R$ and

$$K=\frac{\det(\boldsymbol{\Psi})}{N_R\left[\prod_{k<l}^{N_t}(\alpha_l-\alpha_k)\right]\prod_{l=1}^{N_R-1}l!}.$$

$\boldsymbol{\Psi}$ is the following $\varrho\times\varrho$ Vandermonde matrix

$$\Psi=\begin{bmatrix}1 & \alpha_1 & \cdots & \alpha_1^{\varrho-1}\\ \vdots & \ddots & \ddots & \vdots\\ 1 & \alpha_{\varrho} & \cdots & \alpha_{\varrho}^{\varrho-1}\end{bmatrix}.$$

In addition, $\alpha_1,\ldots,\alpha_{\varrho}$ represent the ϱ distinct eigenvalues of $\bar{\mathbf{R}}_T$. $\mathcal{D}(i,j)$ is the (i,j) cofactor of the $N_R\times N_R$ matrix $\mathbf{C}$ whose (l,k) element equals to

$$[\mathbf{C}]_{l,k}=(k-1)!\left(\alpha_{\varrho+l}^{\varrho+k-1}-\sum_{p=1}^{\varrho}\sum_{q=1}^{\varrho}[\boldsymbol{\Psi}^{-1}]_{p,q}\alpha_{\varrho+l}^{p-1}\alpha_q^{\varrho+k-1}\right).$$

At this point, given that the marginal p.d.f. of $\bar{\lambda}$ is available in (46), we are able to evaluate the following expectation

$$\mathrm{E}_{\bar{\lambda}}\left[\frac{\bar{\lambda}}{(\bar{\lambda}+\vartheta)^2}\right]=\int_0^{+\infty}\frac{\bar{\lambda}}{(\bar{\lambda}+\vartheta)^2}f_{\bar{\lambda}}(x)dx \tag{47}$$

and plug it in (45). Hence, we can prove that:

$$\bar{P}_s=KN_R\sum_{i=1}^{N_R}\sum_{j=1}^{N_R}\mathcal{D}(i,j)\left[\alpha_{\varrho+i}^{\varrho-1}g_j\left(\frac{1}{\alpha_{\varrho+i}}\right)-\sum_{l=1}^{\varrho}\sum_{k=1}^{\varrho}[\boldsymbol{\Psi}^{-1}]_{k,l}\alpha_{\varrho+i}^{k-1}\alpha_l^{\varrho-1}g_j\left(\frac{1}{\alpha_l}\right)\right], \tag{48}$$

where

$$g_j(y)=\vartheta^{j-1}e^{\vartheta y}j!\left[\Gamma(1-j,y\vartheta)-y\vartheta\Gamma(-j,y\vartheta)\right] \tag{49}$$

and $\Gamma(\cdot)$ denotes the gamma function defined in [41, p. 892].

IV. Extension to a Distributed Relays Architecture

In some scenarios there is no backhaul-link between the RNs. This is the case when: i) the delay requirement does not allow communication between the RNs, ii) the cost of installing a backhaul connection is high, or iii) the RNs are remote terminals (like cell phones). This section extends the proposed centralized relay architecture from Section II to a distributed one with the employment of a decentralized detection algorithm at the RNs. All the other configurations of the proposed architecture remain the same.

During the first hop, the received signal at the j-th RN is described in (6). It is straightforward to see that if we set $s_0=0$, (6) can be easily interpreted as the following $(M+1)$-ary Hypothesis Test (HT)

$$H_i: y_{Rj}=d_j s_i+w_{SRj},\quad i=0,\ldots,M, \tag{50}$$

which can be independently employed at each RN. In (50), s_i takes values from $\{s_0,s_1,\ldots,s_M\}$. Clearly, the H_0 hypothesis $(s_0=0)$ corresponds to the case where the j-th RN is not the receiving node in the first hop. Thus, this RN remains silent during the second hop. On the other hand, all the other $H_1,\ldots,H_M$ hypotheses correspond to the case where the j-th node is the receiving/activating RN and the relayed conventional symbol is $s_1,\ldots,s_M$, respectively.

In this case, the distributed ML detector at each RN is reduced to the usual minimum distance rule

$$(\hat{s}_i)=\arg\min_{s_i\in\{s_0,s_1,\ldots,s_M\}}\left|y_{R_j}-d_j s_k\right|. \tag{51}$$

Given that $d_j s_k$ is precomputed and stored in the memory before the beginning of transmission, the computational complexity of (51) at each RN is just one complex subtractions per iteration.

Clearly, the independent execution of the detector of (51) at each RN could result into the activation of multiple RNs during the second hop. In the other extreme case, the distributed detectors of (51) of all RNs could detect the zero symbol s_0. When this extreme case occurs, all the RNs remain silent and no symbol is relayed to the DN. Due to the fact that the DN is unaware of this situation, its detector decides randomly based only on the ever present Gaussian noise.

In the literature, these problems are treated using error-aware detection at the DN [28], [32], which increases the complexity of the DN. In order to keep the complexity of the DN low, we decide to ignore the phenomenon of the activation of multiple RNs. Though, as it is demonstrated in Section VI, our DD algorithm benefits from the linear precoding at the SN, which results in high SNR gains at the RNs. Thus, the probability of activating multiple RNs is reduced. Because of this, practically and in most scenarios, either there is no performance difference from the CD algorithm, or there is a small BER penalty.

V. Theoretical Evaluation of the Average Bit Error Probability

In this section, we employ the well known union bound technique [34] in order to provide bounds on the ABEP of the proposed architecture. This is done both for the CD (Section II) and DD (Section IV) algorithms in the RNs. In order to provide tractability to our theoretical analysis, we focus on the scenario of ZF precoding with P-CSIT at the SN. We make this decision because it is well known that the ABEP performance analysis of MMSE or any other regularized ZF-like precoding method is a challenging problem [43]. Besides, in high SNRs, ZF precoding approaches the performance of MMSE precoding [36], [44].

A. Average Bit Error Probability When the RNs Employ Centralized Detection

When the CD algorithm of (7) is employed at the RNs, the DH-HSSK/HSM is a regenerative relay system with coordinated detection. Thus, we can follow a similar procedure such as [27], [45] and express the end-to-end ABEP as:

$$P_{\mathrm{CD}}(\gamma_1,\gamma_2) = P_1(\gamma_1) + P_2(\gamma_2) - P_1(\gamma_1)P_2(\gamma_2). \quad (52)$$

In (52), γ_1 and γ_2 denotes the SNRs of the first and the second hop, respectively. $P_{\mathrm{CD}}(\gamma_1,\gamma_2)$ is the end-to-end ABEP, given γ_1 and γ_2. Furthermore, $P_1(\gamma_1)$ and $P_2(\gamma_2)$ represents the ABEP of the first and second hop, respectively. It is not difficult to see that $P_1(\gamma_1)$ and $P_2(\gamma_2)$ are the ABEP of R-SM and SM, respectively. The ABEP of SM is widely studied in the literature [10], [11], [46], [47]. However, the ABEP of R-SM is a challenging problem, due to the use of ZF precoding [43].

We begin with the derivation of the ABEP of the first hop (ABEP of R-SM). We have to evaluate the Pairwise Error Probability (PEP) of transmitting $\mathbf{x}$ at the RN and receiving $\hat{\mathbf{x}}$ at the RNs. This is done for all possible combinations of $\mathbf{x},\hat{\mathbf{x}} \in \mathcal{A}_1 = \{\mathbf{e}_i s_k | i \in \{1,\dots,N_R\} \cap k \in \{1,\dots,M\}\}$ and $\mathbf{x} \neq \hat{\mathbf{x}}$. Given that (5) holds and the ML detector selects the symbol $\mathbf{x}$ which minimizes the quantity $\|\mathbf{y}_R - \mathbf{D}\mathbf{x}\|_2^2$, a symbol error occurs when $\mathcal{E} = \{\|\mathbf{y}_R - \mathbf{D}\mathbf{x}\|_2^2 > \|\mathbf{y}_R - \mathbf{D}\hat{\mathbf{x}}\|_2^2\}$. If we set $\mathbf{c} = \mathbf{x} - \hat{\mathbf{x}}$ and elaborate $\mathcal{E}$, it can be shown that $\mathcal{E} = \{-\sum_{i=1}^{N_R} d_i \mathrm{Re}\{c_i^* \mathbf{w}_{SR_i}\} > (\mathbf{c}^H \mathbf{D}\mathbf{c}/2)\}$, where c_i and w_{SR_i} are the elements of $\mathbf{c}$ and $\mathbf{w}_{SR}$, respectively. If we consider that $-\sum_{i=1}^{N_R} d_i \mathrm{Re}\{c_i^* \mathbf{w}_{SR_i}\} \sim \mathcal{N}(0, (\sigma^2_{\mathbf{w}_{SR}}/2)\mathbf{c}^H\mathbf{D}^2\mathbf{c})$, it is shown that the instantaneous PEP conditioned on $\mathbf{D}^2$ is expressed as:

$$P_{\mathrm{R-SM}}(\mathbf{x} \to \hat{\mathbf{x}} | \mathbf{D}^2) = Q\left(\sqrt{\frac{\mathbf{c}^H\mathbf{D}^2\mathbf{c}}{2}\gamma_1}\right), \quad (53)$$

where $\gamma_1 = 1/\sigma^2_{\mathbf{w}_{SR}}$ and $Q(\cdot)$ is the Q-function. We are interested in deriving the PEP by averaging (53) over all possible realizations of $\mathbf{D}^2$:

$$P_{\mathrm{R-SM}}(\mathbf{x} \to \hat{\mathbf{x}}) = \mathrm{E}_{\mathbf{D}^2}\left[Q\left(\sqrt{\frac{\mathbf{c}^H\mathbf{D}^2\mathbf{c}}{2}\gamma_1}\right)\right]. \quad (54)$$

Just like SM, in R-SM there are three possible types of symbol errors: i) signal errors denoted as $\mathcal{E}_1 = \{\mathbf{x} \to \hat{\mathbf{x}} | \{s_k \to \hat{s}_k, \mathbf{e}_i \to \mathbf{e}_i\}\}$; ii) space errors denoted as $\mathcal{E}_2 = \{\mathbf{x} \to \hat{\mathbf{x}} | \{s_k \to s_k, \mathbf{e}_i \to \mathbf{e}_{\hat{i}}\}\}$; and iii) joint signal and space errors denoted as $\mathcal{E}_3 = \{\mathbf{x} \to \hat{\mathbf{x}} | \{s_k \to \hat{s}_k, \mathbf{e}_i \to \mathbf{e}_{\hat{i}}\}\}$. A signal error occurs when the receiving antenna (represented by $\mathbf{e}_i$) is correctly detected, while the transmitted symbol s_k is incorrectly detected. In contrast, a space error takes place when s_k is correctly detected and the receiving antenna is incorrectly detected. Finally, the incorrect detection of s_k and $\mathbf{e}_i$ creates a joint symbol error. Thus, depending on the type of the R-SM symbol error, the RV $z = \mathbf{c}^H\mathbf{D}^2\mathbf{c}$ takes the following form:

$$z = \begin{cases} |s_k - \hat{s}_k|^2 d_i^2, & \text{for } \mathcal{E}_1, \quad (55a) \\ |s_k|^2 \left(d_i^2 + d_{\hat{i}}^2\right), & \text{for } \mathcal{E}_2, \quad (55b) \\ |s_k|^2 d_i^2 + |\hat{s}_k|^2 d_{\hat{i}}^2, & \text{for } \mathcal{E}_3. \quad (55c) \end{cases}$$

As we see from (55a)–(55c), z depends on the RVs d_i^2, $i = 1,\dots,N_R$. Thus, before continuing with the evaluation of (54), let us provide the p.d.f. of d_i^2, $i = 1,\dots,N_R$, which is fundamental for our proof. For ZF precoding, it holds that:

$$d_i^2 = \frac{1}{\left[\left(\mathbf{H}_{SR}\mathbf{H}_{SR}^H\right)^{-1}\right]_{i,i}}. \quad (56)$$

In [48], it is given that $d_i^2 \sim \mathrm{Gamma}(L,1)$, where $L = N_t - N_R + 1$. Here, $\mathrm{Gamma}(k,\theta)$ denotes a gamma distribution with shape k and scale θ. Thus, the RVs d_i^2, $i = 1,\dots,N_R$ have the following p.d.f.:

$$f_{d_i^2}(x) = \frac{1}{\Gamma(L)} x^{L-1} e^{-x} H_0(x). \quad (57)$$

In the following, we evaluate (54) for all different types of R-SM symbol errors. Let us begin with the case of the signal errors. If we set $\delta = s_k - \hat{s}_k$, plug (55a) into (54), and use (57)

in order to average over d_i^2, we can bound the PEP of a signal error as:

$$P_{\mathrm{R-SM}}(\mathbf{x}\to\hat{\mathbf{x}}|\mathcal{E}_1) = \mathrm{E}_{d_i^2}\left[Q\left(\sqrt{\frac{|\delta|^2 d_i^2}{2}\gamma_1}\right)\right] \leq \frac{1}{2}\left[\frac{|\delta|^2}{4}\gamma_1+1\right]^{-L}. \quad (58)$$

In the previous procedure, we use the Chernoff bound $Q(x)\leq (1/2)e^{-(x^2/2)}$ [34]. *We note that (58) is also the PEP of a Multiple-Input Single Output (MISO) system with ZF precoding. As it can be seen, this MISO system has a diversity order of $L=N_t$ (given that there is one receive antenna). This conclusion agrees with the indirect diversity analysis of* [43].

As implied by (55b) and (55c), the evaluation of (54) for the cases of: i) space and ii) joint errors, requires the joint p.d.f. of d_i^2 and $d_{\hat{i}}^2$. From (56), it is clear that d_i^2 and $d_{\hat{i}}^2$ are correlated RVs. Given that d_i^2 and $d_{\hat{i}}^2$ are correlated and distribute as $\mathrm{Gamma}(L,1)$, their joint p.d.f. is the Kibble's bivariate gamma distribution:

$$h_{d_i^2,d_{\hat{i}}^2}(x,y) = \frac{(1-\rho_c)^{-L}}{\Gamma(L)}\sum_{k=0}^{+\infty} b_k f_k(x) p_k(y), \quad (59)$$

as it is given in [49]. Here, we have that:

$$b_k = \frac{\rho_c^k}{(1-\rho_c)^{2k}\Gamma(L+k)k!}, \quad (60)$$

$$f_k(x) = \left[x^{L+k-1}e^{\frac{-x}{1-\rho_c}}\right]H_0(x), \quad (61)$$

and

$$p_k(y) = \left[y^{L+k-1}e^{\frac{-y}{1-\rho_c}}\right]H_0(y). \quad (62)$$

In addition, $\rho_c = \mathrm{E}[(d_i^2-L)(d_{\hat{i}}^2-L)]/L$, which is the Pearson product-moment correlation coefficient. Given the marginal gamma p.d.f. of each d_i and a given system configuration, ρ_c can be robustly estimated.

In order to provide a bound on the PEP when a space error occurs, we use (55b) and average (54) over all realizations of d_i^2 and $d_{\hat{i}}^2$, using (59). Due to space limitation, we do not provide the details. The attained bound is written as:

$$P_{\mathrm{R-SM}}(\mathbf{x}\to\hat{\mathbf{x}}|\mathcal{E}_2) = \mathrm{E}_{d_i^2,d_{\hat{i}}^2}\left[Q\left(\sqrt{\frac{|s_k|^2(d_i^2+d_{\hat{i}}^2)}{2}\gamma_1}\right)\right] \leq \frac{(1-\rho_c)^{-L}}{2t^{2L}}{}_1F_0\left(L;\,;\frac{\rho_c}{(1-\rho_c)^2t^2}\right), \quad (63)$$

where $t=(|s_k|^2\gamma_1/4)+(1/(1-\rho_c))$ and

$${}_1F_0(L;\,;y) = \sum_{k=0}^{+\infty}\frac{[L]_k}{k!}y^k \quad (64)$$

is the generalized hypergeometric function, as defined in [41, p. 1010]. Moreover, $[L]_k = L(L+1)\ldots(L+k-1)$ is the Pochhammer symbol.

Similarly, the bound on the PEP of the joint error of $\mathcal{E}_3$ is achieved by evaluating the expectation of (54), using (55c) and (59):

$$P_{\mathrm{R-SM}}(\mathbf{x}\to\hat{\mathbf{x}}|\mathcal{E}_3) = \mathrm{E}_{d_i^2,d_{\hat{i}}^2}\left[Q\left(\sqrt{\frac{|s_k|^2 d_i^2+|\hat{s}_k|^2 d_{\hat{i}}^2}{2}\gamma_1}\right)\right] \leq \frac{[(1-\rho_c)t_1t_2]^{-L}}{2}\times{}_1F_0\left(L;\,;\frac{\rho_c}{t_1t_2(1-\rho_c)^2}\right), \quad (65)$$

where $t_1=(|s_k|^2\gamma_1/)+(1/(1-\rho_c))$ and $t_2=(|\hat{s}_k|^2\gamma_1/4)+(1/(1-\rho_c))$.

As a final step for our proof, we employ the bounds of (58), (63), and (65), and express the ABEP of the first hop as:

$$P_1(\gamma_1)\leq\frac{1}{Mk_t}\sum_{\mathbf{x}}\sum_{\mathbf{x}\neq\hat{\mathbf{x}}} d(\mathbf{x}\to\hat{\mathbf{x}})P_{\mathrm{R-SM}}(\mathbf{x}\to\hat{\mathbf{x}}), \quad (66)$$

where $d(\mathbf{x}\to\hat{\mathbf{x}})$ denotes the number of different bits between the bit sequences represented by $\mathbf{x}$ and $\hat{\mathbf{x}}$.

In order to complete the derivation of (52), we express the ABEP of the second hop using the results from [46]. Hence, the ABEP of the second hop is bounded as:

$$P_2(\gamma_2)\leq\frac{1}{Mk_t}\sum_{\mathbf{x}}\sum_{\mathbf{x}\neq\hat{\mathbf{x}}} d(\mathbf{x}\to\hat{\mathbf{x}}), P_{\mathrm{SM}}(\mathbf{x}\to\hat{\mathbf{x}}), \quad (67)$$

where

$$P_{\mathrm{SM}}(\mathbf{x}\to\hat{\mathbf{x}}) = \frac{1}{2}\det\left(\mathbf{I}_{\zeta,\zeta}+\gamma_2\frac{\Lambda}{\sqrt{2}}\right). \quad (68)$$

Here, we have that $\zeta = N_D N_R$ and $\Lambda = \mathbf{I}_{N_D,N_D}\otimes\mathbf{c}\mathbf{c}^H$.

B. Average Bit Error Probability When the RNs Employ Distributed Detection

This section provides the ABEP of DH-HSSK/HSM when the RNs utilize the DD of (51). In this type of uncoordinated detection, the methodology that we follow in the previous subsection does not hold, because during the second hop multiple RNs can be active. In this case, the transmitted symbol vector of the second hop is not a legal SM symbol. Thus, (52) is not valid.

In order to overcome this problem, we use the union bound technique and express the end-to-end ABEP of DH-HSM as:

$$P_{\mathrm{D}}(\gamma_1,\gamma_2)\leq\frac{1}{Mk_t}\sum_{\mathbf{x}}\sum_{\mathbf{x}\neq\hat{\mathbf{x}}} d(\mathbf{x}\to\hat{\mathbf{x}})P_{\mathrm{DD}}(\mathbf{x}\to\hat{\mathbf{x}},\gamma_1,\gamma_2), \quad (69)$$

where $P_{\mathrm{DD}}(\mathbf{x}\to\hat{\mathbf{x}},\gamma_1,\gamma_2)$ is the end-to-end PEP of transmitting $\mathbf{x}$ at the SN and detecting erroneously $\hat{\mathbf{x}}\neq\mathbf{x}$ at the

DN, given γ_1 and γ_2 (the SNR of each hop). We remind that $\mathbf{x}, \hat{\mathbf{x}} \in \mathcal{A}_1 = \{\mathbf{e}_i s_k | i \in \{1, \ldots, N_R\} \cap k \in \{1, \ldots, M\}\}$.

It is not difficult to see that, using the total probability theorem, the PEP of an end-to-end symbol error is expressed as:

$$P_{\mathrm{DD}}(\mathbf{x} \to \hat{\mathbf{x}}, \gamma_1, \gamma_2) = \sum_{\tilde{\mathbf{x}} \in \mathcal{A}} P_{\mathrm{SR}}(\mathbf{x} \to \tilde{\mathbf{x}}, \gamma_1) P_{\mathrm{RD}}(\tilde{\mathbf{x}} \to \hat{\mathbf{x}}, \gamma_2), \tag{70}$$

where $\tilde{\mathbf{x}} \in \mathcal{A}$ are all the possible detected/transmitted symbol vectors at the RNs. $P_{\mathrm{SR}}(\mathbf{x} \to \tilde{\mathbf{x}}, \gamma_1)$ is the probability that $\mathbf{x}$ is transmitted at the SN and $\tilde{\mathbf{x}}$ is detected/retransmitted at the RNs. Furthermore, $P_{\mathrm{RD}}(\tilde{\mathbf{x}} \to \hat{\mathbf{x}}, \gamma_2)$ represents the probability that $\tilde{\mathbf{x}}$ is transmitted at the RNs and $\hat{\mathbf{x}}$ is detected at the DN.

Before continuing, we define the set $\mathcal{A} = \{\mathcal{A}_0, \mathcal{A}_1, \ldots, \mathcal{A}_{N_R}\}$. Here, we have that $\mathcal{A}_0 = \{\mathbf{0}_{N_R,1}\}$. The subset $\mathcal{A}_0$ corresponds to the very special case where all the RNs remain silent during the second hop (for more details, the reader is referred to Section IV). Every other subset $\mathcal{A}_q$, with $q = 1, \ldots, N_R$ and $|\mathcal{A}_q| = \binom{N_R}{q} M^q$, contains all the possible vectors, of size $N_R \times 1$, with exactly q non zero elements. Their non zero elements take values from $\mathcal{M} = \{s_1, \ldots, s_k\}$. Mathematically, this is represented as $\mathcal{A}_q = \{\mathbf{e}_{i_1} s_{k_1} + \cdots + \mathbf{e}_{i_q} s_{k_q} | \mathcal{B}_1 \cap \mathcal{B}_2\}$, where $\mathcal{B}_1 = \{\{i_1, \ldots, i_q\} \in \{1, \ldots, N_R\} | \{i_1 \neq \cdots \neq i_q\}\}$, $\mathcal{B}_2 = \{\{k_1, \ldots, k_q\} \in \{1, \ldots, M\}\}$, and $\mathbf{e}_{i_l}$ is the i_l-th column of $\mathbf{I}_{N_R, N_R}$. We note that, clearly, the correct symbol $\mathbf{x}$ belongs to $\mathcal{A}$. Even thought the RNs might be able to detect correctly the transmitted vector $\mathbf{x}$, a symbol error could occur during the second hop.

We have to evaluate each probability of the right hand side of (70). We begin with $P_{\mathrm{SR}}(\mathbf{x} \to \tilde{\mathbf{x}}, \gamma_1)$. Given that detection is conducted independently at each RN and the Gaussian noise is independent, we have:

$$P_{\mathrm{SR}}(\mathbf{x} \to \tilde{\mathbf{x}}, \gamma_1) = \prod_{j=1}^{N_R} \mathrm{E}_{d_j^2} \left[P_{\mathrm{SR}_j} \left(\mathbf{x}_j \to \tilde{\mathbf{x}}_j, \gamma_1 | d_j^2 \right) \right]. \tag{71}$$

Here, $\mathbf{x}_j, \tilde{\mathbf{x}}_j \in \{s_0, \mathcal{M}\}$ are the j-th elements of $\mathbf{x}$ and $\tilde{\mathbf{x}}$, respectively. In simple words, $\tilde{\mathbf{x}}_j$ belongs to the employed constellation $\mathcal{M}$ or is 0 and $\tilde{\mathbf{x}}_j$ is the detected symbol which again belongs to $\mathcal{M}$ or is zero.

Thus, we have to evaluate

$$P_{\mathrm{SR_j}}(\mathbf{x}_j \to \tilde{\mathbf{x}}_j, \gamma_1) = \mathrm{E}_{d_j^2} \left[P_{\mathrm{SR}_j} \left(\mathbf{x}_j \to \tilde{\mathbf{x}}_j, \gamma_1 | d_j^2 \right) \right] \tag{72}$$

by integrating over all the possible realizations d_j^2. The received signal at the j-th RN is given from (6). Hence, given that $\mathbf{x}_j, \tilde{\mathbf{x}}_j \in \{s_0, \mathcal{M}\}$, with $\mathbf{x}_j \neq \tilde{\mathbf{x}}_j$, we can write

$$P_{\mathrm{SR}_j} \left(\mathbf{x}_j \to \tilde{\mathbf{x}}_j, \gamma_1 | d_j^2 \right) = Q \left(\sqrt{\frac{\gamma_1 |\mathbf{x}_j - \tilde{\mathbf{x}}_j|^2 d_j^2}{2}} \right). \tag{73}$$

For this reason, if we follow the same steps as the derivation of (58), we can prove that:

$$P_{\mathrm{SR}_j} \left(\mathbf{x}_j \to \tilde{\mathbf{x}}_j, \gamma_1 \right) \leq \frac{1}{2} \left[\frac{|\mathbf{x}_j - \tilde{\mathbf{x}}_j|^2}{4} \gamma_1 + 1 \right]^{-L}. \tag{74}$$

For the case of $\mathbf{x}_j = \tilde{\mathbf{x}}_j$ (correct transmission to the j-th RN), we have that

$$\begin{aligned} P_{\mathrm{SR}_j} \left(\mathbf{x}_j \to \tilde{\mathbf{x}}_j, \gamma_1 \right) &= 1 - P_e(\gamma_1) \\ &\leq 1 - \sum_{\substack{j=0 \\ \mathbf{x}_j \neq s_j}}^{M} P_{\mathrm{SR}_j}(\mathbf{x}_j \to s_j, \gamma_1), \end{aligned} \tag{75}$$

where $P_e(\gamma_1)$ is the probability of erroneous detection. Hence, the probability of interest is bounded by

$$P_{\mathrm{SR}_j}(\mathbf{x}_j \to \tilde{\mathbf{x}}_j, \gamma_1) \leq 1 - \frac{1}{2} \sum_{\substack{j=0 \\ \mathbf{x}_j \neq s_j}}^{M} \left[\frac{|\mathbf{x}_j - s_j|^2}{4} \gamma_1 + 1 \right]^{-L}. \tag{76}$$

In addition, the evaluation of (70) requires the knowledge of $P_{\mathrm{RD}}(\tilde{\mathbf{x}} \to \hat{\mathbf{x}}, \gamma_2)$. We know that $\tilde{\mathbf{x}} \in \mathcal{A}$, which means that the RNs retransmit legal or illegal SM symbols. Even in the case of the retransmission of illegal SM symbols, as we explain in Section IV, the DN uses the detector of (9). Thus an error happens when $\mathcal{E}' = \{\|\mathbf{y}_D - \mathbf{H}_{RD}\mathbf{x}\|_2^2 > \|\mathbf{y}_D - \mathbf{H}_{RD}\hat{\mathbf{x}}\|_2^2 | \tilde{\mathbf{x}}\}$ (given that $\tilde{\mathbf{x}}$ is relayed by the RNs). We can easily transform $\mathcal{E}'$ as $\mathcal{E}' = \{\|\mathbf{H}_{RD}\mathbf{c}_1 + \mathbf{w}_{RD}\|_2^2 > \|\mathbf{H}_{RD}\mathbf{c}_2 + \mathbf{w}_{RD}\|_2^2\}$, where $\mathbf{c}_1 = \tilde{\mathbf{x}} - \mathbf{x}$ and $\mathbf{c}_2 = \tilde{\mathbf{x}} - \hat{\mathbf{x}}$. We can further proceed and reach to $\mathcal{E}' = \{\|\mathbf{H}_{RD}\mathbf{c}_1\|_2^2 - \|\mathbf{H}_{RD}\mathbf{c}_2\|_2 > 2\mathrm{Re}\{\mathbf{c}^H \mathbf{H}_{RD}^H \mathbf{w}_{RD}\}\}$, where $\mathbf{c} = \mathbf{c}_2 - \mathbf{c}_1$. Given that $-\sum_{k=1}^{ND} \mathrm{Re}\{\mathbf{w}_{RD_k} a_k\} \sim \mathcal{N}(0, (\sigma_{w_{RD}}^2/2)\|\mathbf{H}_{RD}\mathbf{c}\|_2^2)$, where $a_k = \sum_{j=1}^{N_D} \mathbf{c}_j^* [\mathbf{H}_{RD}]_{j,k}^*$, we have that the instantaneous probability for which we are interested is written as:

$$P_{\mathrm{RD}}(\tilde{\mathbf{x}} \to \hat{\mathbf{x}}, \gamma_2 | \mathbf{H}_{RD}) = Q \left(\gamma' \frac{\|\mathbf{H}_{RD}\mathbf{c}_2\|_2^2 - \|\mathbf{H}_{RD}\mathbf{c}_1\|_2^2}{\sqrt{\|\mathbf{H}_{RD}\mathbf{c}\|_2^2}} \right), \tag{77}$$

where $\gamma' = \sqrt{\gamma_2/2}$. The direct evaluation of the expectation of (77), over all channel realizations, is a difficult task. Actually, this expectation requires the p.d.f. of the summation of two Erlang (gamma) RVs divided by a generalized gamma RV, where all RVs are correlated. An approach that gives a tractable solution to the expectation of (77) is to use the triangle and inverse triangle inequality and upper and lower bound the RV $Y = (\|\mathbf{H}_{RD}\mathbf{c}_2\|_2^2 - \|\mathbf{H}_{RD}\mathbf{c}_1\|_2^2)/\sqrt{\|\mathbf{H}_{RD}\mathbf{c}\|_2^2}$. Unfortunately, this method gives bounds that are loose and do not provide any insight. Due to space limitation, we do not include this analysis. An alternative way to evaluate the expectation of (77) is numerically. In this way, we conclude our proof.

VI. Simulation Results and Discussion

In this section, we present theoretical and Monte Carlo simulation results that demonstrate the performance of the proposed DH-HSSK/HSM architecture under different forms of CSIT at the SN, both for the CD and DD algorithms. We consider scenarios with: i) small, ii) medium and iii) large system configuration.

For the sake of comparison, we compare our architecture against four benchmark systems: i) the corresponding single

relay system, ii) the best relay selection system, iii) a form of distributed Alamouti relaying, and iv) the direct communication link. We note that in all the benchmark systems, the operating principle of the SN and the DN is the same as DH-HSM/HSSK. The SN uses the same linear precoding methods and the DN detects the received signal using ML detection.

The corresponding single relay system is denoted as Dual Hop-with one Relay (DH-w1R). The adoption of this form of comparison for the evaluation of space modulated VMIMO is extensively employed in the literature [20], [24], [27], [28], [32]. The operational principle of DH-W1R is similar to the proposed DH-HSSK/HSM architecture with the following exception. The conveyed constellation point is relayed by a single RN.

We note that the RN of DH-w1R detects the relayed constellation point using a ML detector which has almost the same complexity as the ML detector of (7) and higher complexity than the detector of (51) (all detectors are single stream detectors). Especially, the DD of (51) results in a significant lower complexity as the spectral efficiency and the number of the RNs is increased. Because of this, DH-HSSK/HSM can be interpreted as a way of distributing the computational complexity of the single RN of DH-W1R to the multiple RNs of DH-HSSK/HSM. An additional advantage of DH-HSSK/HSM compared with DH-W1R is that, under the assumption of high SNR at the RNs, usually only one out of N_R RNs is retransmitting the symbol received from the SN. This is very important when the RNs are remote terminals (for example mobile phones) with a battery supply, because a battery life extension is achieved.[1] Hence, if we take into account the advantages of DH-HSSK/HSM compared with DH-W1R with respect to: i) the complexity and power consumption distribution and ii) the BER gains (as shown in the following), it is clear that the cost of having multiple RNs of low complexity and extended battery life compared with the cost of having a single RN of high complexity is insignificant.

The next two benchmark systems that we employ is the: i) best relay selection technique denoted as Dual Hop-with Best Relay Selection (DH-wBRS) and ii) the distributed Alamouti code represented as Dual Hop-Distributed Alamouti Code (DH-DAC). In DH-wBRS, a single RN out of N_R is selected to convey the transmitted constellation point in every block of symbols. DH-wBRS selects the RN with the most favorable RN-DN channel which has the highest channel gain. The selection of this RN offers the highest instantaneous receive SNR at the DN. In DH-DAC, two symbols are conveyed to two RNs in order to form a distributed Alamouti codeword. The RNs can employ a CD or DD algorithm, just like DH-HSM. In the DD case, the orthogonality of the Alamouti codeword cannot be ensured, whereas in the CD case it can be ensured. We note that DH-DAC requires 4 symbol period in order to convey two constellation points (constructed Alamouti codeword) from the SN to the DN. Clearly, the comparison of DH-HSM with DH-wBRS and DH-DAC is not entirely fair. DH-wBRS uses closed loop transmission during the second hop, while DH-HSM uses open loop transmission. Furthermore, DH-DAC activates two RNs during the second hop, whereas DH-HSM activate only one. Finally, we state that: i) DH-DAC is used as a benchmark system only when the RNs are two and ii) no best RN selection criterion is applied in DH-DAC. As we see in the following, even though these comparisons are not entirely fair, DH-HSM is able to achieve significant performance gains in the majority of the cases.

The last benchmark system that we use is the direct SN-DN MIMO communication. We assume that the SN-DN wireless channel is weak and distributes as $\mathcal{CN}(\mathbf{0}, p\mathbf{I})$, where p captures the path loss effect. This assumption is valid because DH-HSM is designed to be applied when the SN-DN channel is weak.

For all systems, the SNR during the first hop is defined as $\gamma_1 = \text{SNR}_{\text{dB}}^{\text{SR}} = 10\log_{10}(P_s^{SR}/\sigma_{w_{SR}}^2)$, where $P_s^{SR} = \text{E}_{\mathbf{x}}[\|\mathbf{PDx}\|_2^2]$ is the power transmitted at SN. For a fair comparison between DH-HSSK/HSM and DH-w1R, we set $P_s = 1$ by using the appropriate: i) normalization matrix $\mathbf{D}$ for the precoder $\mathbf{P}$ (as described in Section II) and ii) the appropriate normalization of the conventional transmitted constellation diagram. During the second hop, the SNR is defined as $\gamma_1 = \text{SNR}_{\text{dB}}^{\text{RD}} = 10\log_{10}(P_s^{RD}/\sigma_{w_{RD}}^2)$, where $P_s^{RD} = 1$ is the transmitted power from the active RN. Without loss of generality we assume that $\gamma_1 = \gamma_2$. Finally, we assume that the SN–RNs and RNs–DN wireless channel remain constant over a block symbol. In addition, when I-CSIT is available at the SN, $\bar{\mathbf{H}}_{\text{SR}}$ does not change over a block of symbols.

A. Confirmation of the Average Bit Error Probability of DH-HSM Using Simulation Results

Before we proceed with the comparison of DH-HSM with the benchmark system, in Fig. 2, we demonstrate how close are the obtained simulation results with the theoretical framework of Section V. Fig. 2(a) presents the bound of ABEP when CD is employed at the RNs. In addition, Fig. 2(b) shows the ABEP bound under DD at the RNs. We consider two system configurations for CD. In the first configuration, we have a $8 \times 4 \times 4$ ($N_t = 8$, $N_R = 4$, and $N_D = 4$) system with $k_t = 1$ bpsp for DH-HSSK and $k_t = 2$ bpsp for DH-HSM. Furthermore, the second configuration is a $16 \times 8 \times 4$ system with $k_t = 1.5$ bpsp for DH-HSSK and $k_t = 2.5$ bpsp for DH-HSM. As can be seen from Fig. 2(a), for CD and in high SNR, the theoretical and simulation curves perfectly match. In low SNR, there is a difference which is a well known phenomenon caused by the union bound technique [34]. In addition, Fig. 2(b), demonstrates that the bound obtained for DD in Section V-B is close to the simulated curves.

B. Small Scale System Setup

Fig. 3 explores the BER performance of DH-HSM when P-CSIT is available at the DN and the system setup is $4 \times 2 \times 4$, with $k_t = 4$ bpsp. For the sake of comparison, in Fig. 3, we include the performance of: i) DH-w1R, ii) DH-wBRS, iii) DH-DAC, and iii) direct link communication when $p = \{0.1, 1\}$. Especially, the scenario of the direct link communication with

[1] In this paper, we assume that the energy consumption for the RF power transmission is relatively high compared to the energy consumption of the circuits of the RNs during the detection period.

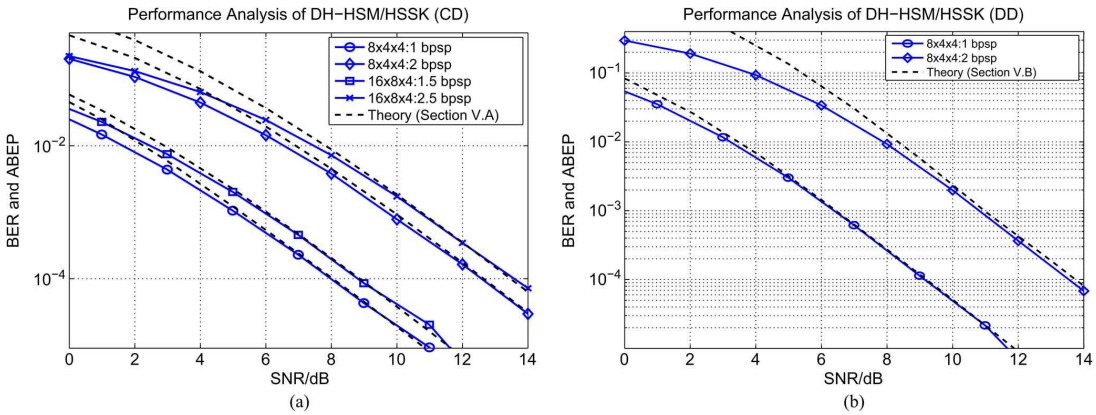


Fig. 2. Performance analysis of DH-HSM when ZF with P-CSIT is employed at the RNs, using: i) simulation results and ii) the bounds of Section V. (a) The RNs employ the CD algorithm. (b) The RNs employ the DD algorithm.

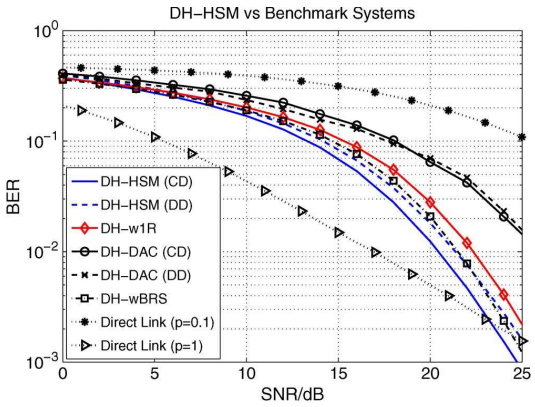


Fig. 3. BER performance of a small scale DH-HSM $(4 \times 2 \times 4)$ system versus DH-w1R, DH-wBRS, DH-DAC, and the 4×4 direct link communication when $p = \{0.1, 1\}$. The spectral efficiency is $k_t = 4$ bpsp.

a weak SN-DN channel $(p = 0.1)$ is a typical application scenario of DH-HSM (or any other dual hop system). Furthermore, as a reference point, we include the performance of direct link communication when the SN-DN channel is strong with $p = 1$ (same statistics as the SN–RNs and RNs–DN channels).

As it can be seen from Fig. 3, DH-HSM offers better BER performance than DH-w1R, DH-wBRS, and DH-DAC due to its multiplexing gain. It quite notable that in the depicted SNR region, DH-HSM offers better performance than DH-wBRS and DH-DAC. Even though, this comparison is not entirely fair. We note that as SNR increases to extremely high value, DH-wBRS and DH-DAC would offer better performance, due to their diversity achieving techniques employed during the second hop. We note that, just like DH-HSM, the performance of DH-DAC is not affected by the employment of CD or DD at the RNs. Regarding the performance of the direct link communication, we see that when the SN-DN channel is weak $(p = 0.1)$, its performance is significantly worse (as expected). For a strong SN-DN channel $(p = 1)$, we see that the direct link offers better performance due to its multiplexing gain. Though, in high SNR, the direct link is remarkably outperformed due to its unity diversity order [43].

C. Medium Scale Setup: The Effect of Multiple Antennas at the DN

The performance evaluation of DH-HSSK/HSM is presented in Fig. 4. In the same figure, we depict the performance of DH-w1R and DH-wBRS for the purpose of comparison. We note that in the following results, we do not use the direct link communication and distributed STBC as benchmark systems. We do not employ the direct link communication, because we propose DH-HSM for a scenario where either the SN-DN is very weak (where direct link communication is clearly outperformed), or there is no SN-DN link. Moreover, distributed STBC is not further used, because as the size of the distributed STBC codeword increases, its fractional STBC rate determines performance significantly worse than DH-HSM.

In addition, we note that for DH-W1R and DH-wBRS, the BER performance is depicted only for ZF precoding at the SN. It can be proved that in the very special case of a MISO linear precoding system (first hop) and when the normalization process of Section II is applied, the MMSE and the ZF forcing precoders result in the same receiving signal. Hence, the performance is the same.

Fig. 4(a) and (b) demonstrate that when the number of receive antennas is small $N_D = \{1, 2\}$, there is no performance difference between the CD and DD algorithms of DH-HSSK/HSM. As N_D is increased, there is a difference. For $N_D = 4$, the difference can be noticed only in low SNRs. Though, when $N_D = 16$, this performance gap is significantly increased. This performance difference exists due to the lost multiplexing gain which is caused by the activation of multiple RNs.

Regarding the comparison with the benchmark systems, we see that for $N_D > 1$, DH-HSSK/HSM is significantly better than DH-w1R. When, $N_D = 1$, DH-HSSK offer a slightly better performance than DH-w1R. Moreover, DH-HSM is

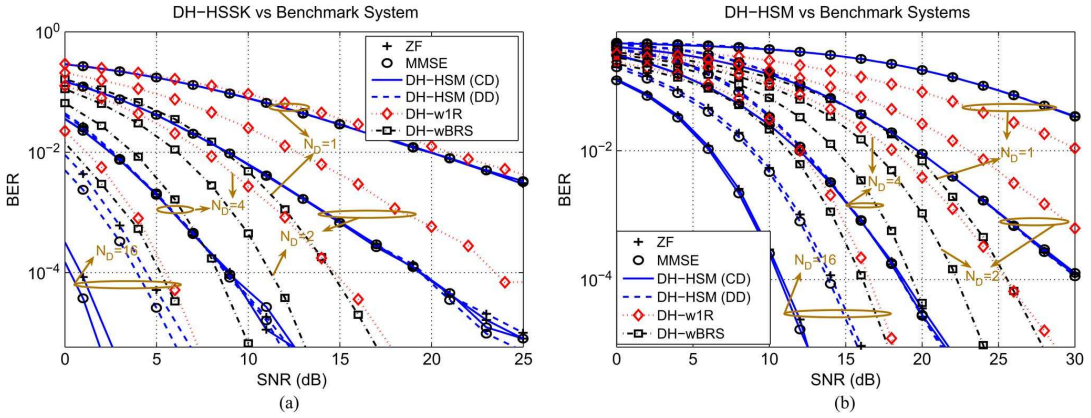


Fig. 4. BER performance of a medium scale DH-HSSK/HSM system, for $N_D = \{1, 2, 4, 16\}$, when: i) the SN employs P-CSIT and ii) the SNR per hop is the same. (a) $16 \times 8 \times N_D$ DH-HSSK system ($k_t = 1.5$ bpsp). (b) $16 \times 8 \times N_D$ DH-HSM system ($k_t = 3.5$ bpsp).

out-performed due to the unity diversity order of SM during the second hop [50].

In addition, Fig. 4(a) and (b) include the comparison of DH-HSSK/HSM with DH-wBRS. We recall that this comparison is not fair for DH-HSSK/HSM. For $N_D = \{1, 2\}$, DH-wBRS demonstrates better performance. In contrast, for a value of $N_D = 4$ and low SNRs, DH-HSSK/HSM outperforms DH-wBRS due to its higher multiplexing gain. Though, after a certain point ($\sim$7.5 dB and $\sim$20 dB for DH-HSSK and DH-HSM, respectively), DH-wBRS has better performance due to its higher diversity order. Remarkably, when $N_D = 16$ and for the depicted practical BER, DH-HSSK and DH-HSM offer a significantly better performance, even for DD.

D. A Very Large Scale Setup

Recently, the concept of very large scale MIMO (or Massive MIMO) has attracted a great attention from the research community [51]. Thus, in Fig. 5, we explore the performance of DH-HSSK/HSM in a system with a large number of antennas at the SN and DN, and a large number of RNs $(64 \times 32 \times 64)$.

The first conclusion drawn from Fig. 5 is that both DH-HHSK and DH-HSM achieve practical BERs in very low SNRs. For example, DH-HSSK achieves a BER of 10^{-4} at about -5.4 dB when CD is employed and at about -2.2 dB when DD is employed. This phenomenon is attributed to the massive MIMO setup of the first hop [51] and to the massive multiplexing gain and receive diversity obtained from SM during the second hop. In addition, Fig. 5 demonstrates that the employment of ZF or MMSE precoding at the SN results in a performance gap for both DH-HSSK and DH-HSM architectures, and for both CD and DD. Regarding the use of CD or DD at the RNs, we observe that there is a penalty for both DH-HSSK and DH-HSM when DD is employed. This penalty is decreased as the value of SNR is decreased.

Furthermore, Fig. 5 shows that DH-HSSK ($k_t = 2.5$ bpsp) and DH-HSM ($k_t = 4.5$ bpsp), both for CD and DD, have a significantly better performance than DH-w1R and DH-wBRS,

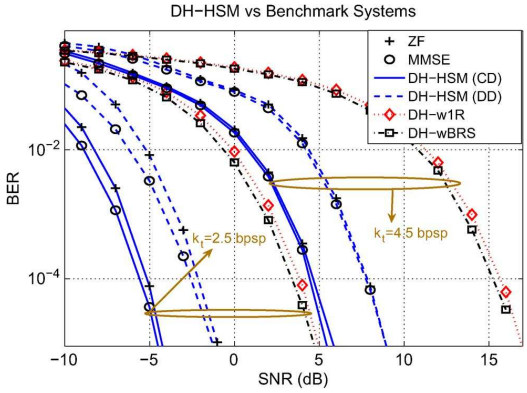


Fig. 5. BER performance of a very large scale DH-HSSK/HSM $(64 \times 32 \times 64)$ system when: i) the SN employs P-CSIT, and ii) the SNR per hop is the same.

for the depicted practical values of BER. We remind the reader that the comparison of DH-HSSK/HSM and DH-wBRS is not entirely fair, due to the closed loop transmission of the second hop of DH-wBRS. It is quite notable, that for BER $= 10^{-4}$, DH-HSSK has a performance difference of about 8 dB when CD is employed and 5 dB when DD is employed at the RNS, compared with DH-wBRS. The performance difference against DH-w1R is further increased to about 9 dB when CD is employed and 6 dB when DD is employed. Additionally, for the same BER, DH-HSM demonstrates a performance gap of about 9 dB for CD and about 6 dB for DD, compared with DH-wBRS. However, when DH-HSM is compared with DH-w1R, the performance difference is further increased to about 10.5 dB for CD and to 7.5 dB for DD. We note that all systems are able to achieve a huge receive diversity gain due to the big number of receive antennas at the DN.

At this point, we note that due to the simplicity, low complexity, and very good BER performance of DH-HSSK/HSM at very low SNRs, our architecture would be a perfect candidate

TABLE I
RELATIVE ENERGY EFFICIENCY GAINS OF DH-HSSK/HSM VERSUS DH-W1R FOR DIFFERENT SYSTEM CONFIGURATIONS

Configuration		k (bpsp)	Target BER	RAER [%]
DH-HSSK	$8 \times 4 \times 1$	1	10^{-2}	0%
with CD	$16 \times 8 \times 4$	1.5	10^{-3}	77.6%
at the RNs	$64 \times 32 \times 64$	2.5	10^{-4}	87%
DH-HSSK	$8 \times 4 \times 1$	1	10^{-2}	0%
with DD	$16 \times 8 \times 4$	1.5	10^{-3}	77.6%
at the RNs	$64 \times 32 \times 64$	2.5	10^{-4}	75%
DH-HSM	$8 \times 4 \times 1$	3	10^{-1}	-14.5%
with CD	$16 \times 8 \times 4$	3.5	10^{-3}	77.6%
at the RNs	$64 \times 32 \times 64$	4.5	10^{-4}	91%
DH-HSM	$8 \times 4 \times 1$	3	10^{-1}	-14.5%
with DD	$16 \times 8 \times 4$	3.5	10^{-3}	77.6%
at the RNs	$64 \times 32 \times 64$	4.5	10^{-4}	82.2%

for wireless backhaul connection between BSs with a very large number of antennas. Such an application could be easily implemented using either fixed or mobile RNs, or a combination of them.

E. RF Energy Efficiency Gains of DH-HSSK/HSM

In order to quantify the RF energy efficiency gain of DH-HSSK/HSM over the benchmark system of DH-W1R, we employ the metric of the Relative Average Energy Reduction (RAER) defined as:

$$\text{RAER}[\%] = \left[1 - 10^{-\frac{\Delta_{\text{SNR}}}{10}}\right] \times 100\%. \tag{78}$$

Δ_{SNR} denotes the SNR (in dB) difference between DH-HSSK/HSM and DH-W1R for a given target BER.

Table I demonstrates the RAER benefits of DH-HSSK/HSM over DH-W1R, for different system configurations and spectral efficiency (in bpsp). In order to conduct a fair comparison between DH-HSSK/DH-HSM and the benchmark system, we use the same: i) BER target and ii) spectral efficiency, for each system configuration. As it can be seen from Table I, in almost all studied scenarios, our DH-HSSK and DH-HSM architectures are more energy efficient than DH-W1R. There are only two exceptions, when the DN is equipped with one receive antenna ($N_D = 1$). In all other scenarios, DH-HSSK and DH-HSM achieve significant energy efficiency gains (in terms of RAER) over DH-W1R. As we see, these gains can be as high as 91%.

F. Power Transmission at the SN and Residual MSE at RNs When I-CSIT Is Available at the SN

In Section III-F, we demonstrate that the unnormalized MMSE, SRA, and WCRA precoders result in a constrained power transmission, while the ZF precoder does not. Hence, when MMSE, SRA, and WCRA precoding designs are employed in a real system implementation, it is possible to avoid the use of the precoding matrix $\mathbf{D}$ at SN. In this case, given the assumption of $\mathbf{H}_{\text{SR}}\mathbf{P} \approx \mathbf{I}_{N_R,N_R}$, there is no need for the i-th RN, $i = 1, \ldots, N_R$ to have the knowledge of d_i. Thus, we are able to employ a non-coherent detection algorithm at the RNs.

In Fig. 6, using simulation results and the theoretical framework of Section III-G, we evaluate the average precoding power $\bar{P}_{\text{s}}$ at the SN and the residual Mean Square Error (MSE) at each of the RNs, with respect to the variance of channel uncertainty $\sigma^2_{\widetilde{\mathbf{H}}_{\text{SR}}}$. This is done for the MMSE, SRA, and WCRA precoding methods when the normalization matrix $\mathbf{D}$ is not applied. We note that due to the fact that $\text{E}_{\mathbf{x}}[\|\mathbf{x}\|_2^2] = 1$, the instantaneous precoding power $P_{\text{s}} = \|\mathbf{P}\|_F^2$ dictates the total transmission power.

As shown in Fig. 6, when $\sigma^2_{\widetilde{\mathbf{H}}_{\text{SR}}}$ is increased, the power of SRA and WCRA precoding methods is decreased. Thus, as the CSIT at the SN becomes more inaccurate, the power transmission is reduced in order to avoid further degradation (inaccurate beamforming). In contrast, when MMSE precoding is employed and for all the depicted values of $\text{SNR}_{\text{SR}} = \{0, 20\}$ dB, the precoding power is increased as the variance of channel uncertainty is increased. This results to more inaccurate beamforming as shown in Fig. 6(b).

Furthermore, Fig. 6 shows that correlation affects the transmitted power of our precoding designs in a composite way. When $\sigma^2_{\widetilde{\mathbf{H}}_{\text{SR}}} \to 0$, higher correlation (higher values of ρ) results in higher precoding power. In contrast, when $\sigma^2_{\widetilde{\mathbf{H}}_{\text{SR}}}$ is increased, higher correlation results in lower precoding power. In contrast to our designs, the existence of channel correlation, when MMSE precoding is used, imposes higher precoding power. As a final remark, we note that all the theoretical curves perfectly match with the simulated curves. In addition, all the claims of Section III-F concerning the relation between the proposed precoding designs and MMSE precoding are confirmed.

Fig. 6(b) demonstrates the residual MSE per RN, defined as $\text{MSE} = \text{E}_{\mathbf{H}_{\text{SR}}}[\|\mathbf{H}_{\text{SR}}\mathbf{P} - \mathbf{I}_{N_R,N_R}\|_F^2]/N_R$ ($\mathbf{I}_{N_R,N_R}$ is the ideal ZF impulse response), which quantifies the residual ICI caused by the inaccurate beamforming, versus the variance of channel uncertainty. As it can be seen, for values of $\sigma^2_{\widetilde{\mathbf{H}}_{\text{SR}}}$ close to zero (almost no uncertainty), MMSE results to lower MSE. It is well known that the MMSE precoder is the optimum linear precoder in terms of MSE. Though, as the SN–RNs channel becomes more inaccurate (higher values of $\sigma^2_{\widetilde{\mathbf{H}}_{\text{SR}}}$) to the SN, our precoding designs offer significantly lower residual MSE than MMSE (for all SNR_{SR} values). In the same figure, we depict the performance of the studied precoding designs when the SN–RNs channel is correlated. Clearly, higher correlation results in higher MSE. Finally, we see that as the channel uncertainty becomes high, the effect of correlation to SRA and WCRA is insignificant compared to the uncorrelated case.

G. BER Performance of DH-HSSK/HSM: The Effect of I-CSIT at the SN

In order to acquire a complete picture of the performance of DH-HSSK/HSM under: i) the effect of I-CSIT and ii) different types of precoding at the SN, in Fig. 7, we present the BER curves for both systems when the variance of the SN–RNs channel uncertainty is $\sigma^2_{\widetilde{\mathbf{H}}_{\text{SR}}} = \{0.01, 0.2\}$. For the sake of comparison, in the same figures we include the performance of DH-HSSK/HSM when MMSE precoding with P-CSIT is

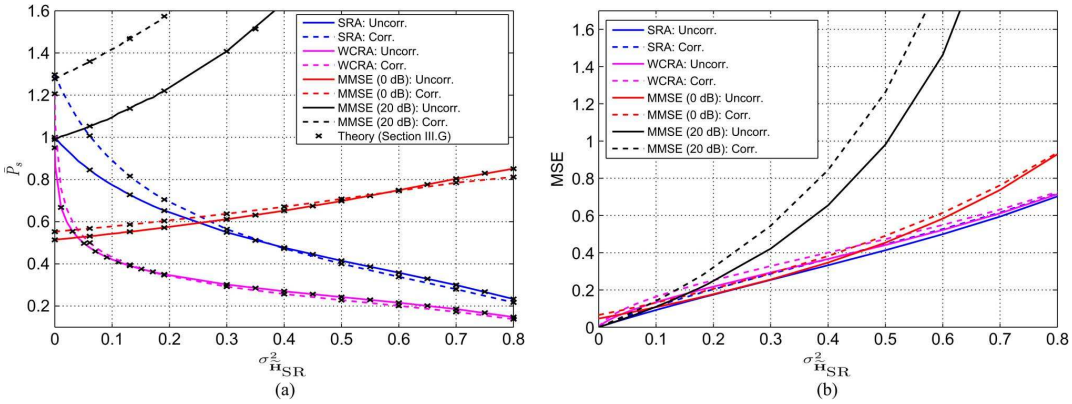


Fig. 6. Effect of I-CSIT at the SN of a $8 \times 4 \times 4$ system. (a) Average precoding power at the SN. (b) Residual MSE at each RN.

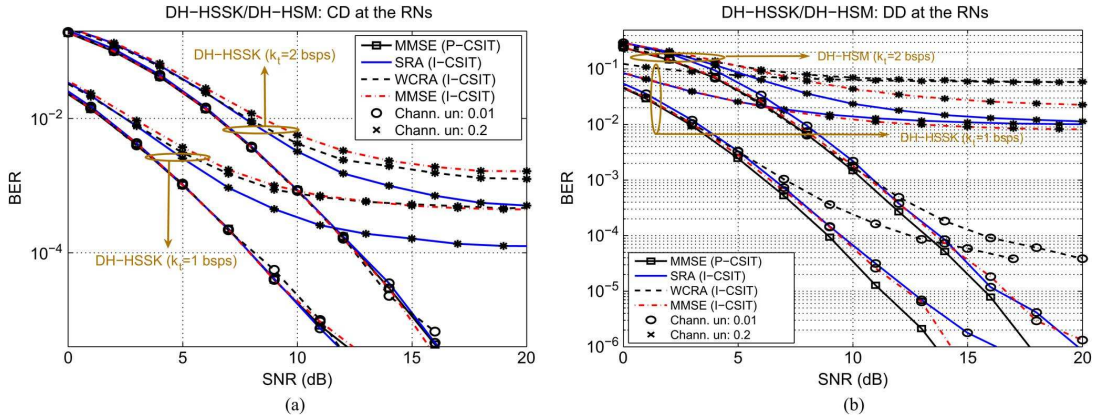


Fig. 7. BER performance of a $8 \times 4 \times 4$ DH-HSSK ($k_t = 1$ bpsp) and a $8 \times 4 \times 4$ DH-HSM ($k_t = 2$ bpsp) system, when different values of channel uncertainty for the SN–RNs channel ($\sigma^2_{\widetilde{\mathbf{H}}_{\mathrm{SR}}} = \{0.01, 0.2\}$) are used. (a) CD at the RNs. (b) DD at the RNs.

employed at the SN. In addition, we assume that the SN–RNs and RNs–DN channels are uncorrelated.[2]

As it can be seen from Fig. 7(a), when CD is employed at the RNs and the channel uncertainty is small ($\sigma^2_{\widetilde{\mathbf{H}}_{\mathrm{SR}}} = 0.01$), DH-HSSK and DH-HSM have no performance difference from the ideal scenario of P-CSIT at the SN. This is true for all types of precoding. In contrast, the increase of channel uncertainty to $\sigma^2_{\widetilde{\mathbf{H}}_{\mathrm{SR}}} = 0.2$ causes a BER performance degradation. In low SNR, the degradation is about 1 dB. Though, an error floor is caused in high SNR, due to the noise limited detection at the RNs. In fact, this is the effect of inaccurate beamforming from the SN. Finally, from Fig. 7(a), we conclude that when I-CSIT is available at the SN, the optimal precoding method is SRA.

In addition, Fig. 7(b) demonstrates that the effect of I-CSIT has a more diminishing result for the case of DD. In fact, in low SNR, there is a small BER performance loss even for low channel uncertainty ($\sigma^2_{\widetilde{\mathbf{H}}_{\mathrm{SR}}} = 0.01$), both for DH-HSSK and DH-HSM. Moreover, after 10 dB, the phenomenon of the error floor in high SNR starts to appear. Finally, we see that high channel uncertainty ($\sigma^2_{\widetilde{\mathbf{H}}_{\mathrm{SR}}} = 0.2$) causes a diminishing BER performance, even in high SNRs.

In Fig. 8, we demonstrate the BER performance of DH-HSSK/DH-HSM under the composite effect of: i) I-CSIT at the SN and ii) correlated SN–RNs and RNs–DN channel. The variance of channel uncertainty at the SN is $\sigma^2_{\widetilde{\mathbf{H}}_{\mathrm{SR}}} = 0.05$. Furthermore, the correlation scenarios that we consider are: i) low correlation with $\rho = 0.1$ and ii) high correlation with $\rho = 0.5$. In addition, given that the RNs are placed far apart, the SN–RNs channel contains only transmit space correlation, while the RNs–DN channel is affected only by receive space correlation. As a benchmark reference, we employ the ideal scenario of P-CSIT at the SN and uncorrelated SN–RNs and RNs–DN channels.

Fig. 8(a) shows that for the depicted SNRs and when correlation is low $\rho = 0.1$, the performance of the proposed architecture with CD does not deviate from the ideal case of no

[2]We note that, in order to provide a fair comparison between the different precoding methods, we make use of the normalization matrix $\mathbf{D}$ as defined in Section II.

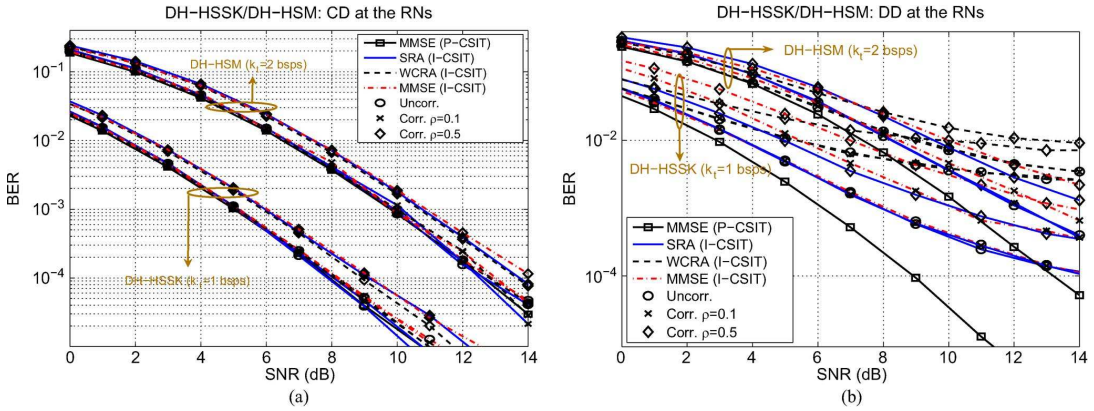


Fig. 8. BER performance of a $8 \times 4 \times 4$ DH-HSSK ($k = 1$ bpsp) and a $8 \times 4 \times 4$ DH-HSM ($k = 2$ bsps) system when: i) both the SN–RNs and the RNs–DN channels are correlated with $\rho = \{0.1, 0.5\}$ and ii) the SN possesses I-CSIT with $\sigma^2_{\tilde{\mathbf{H}}_{\mathrm{SR}}} = 0.05$. (a) CD at the RNs. (b) DD at the RNs.

correlation and P-CSIT at the SN. Furthermore the inspection of the same figure shows that high correlation degrades the BER performance for about 1 dB at BER $= 10^{-3}$. In fact, the previous conclusions are practically true for all methods of precoding. Moreover, Fig. 8(b) presents that the use of DD at the RNs results in a diminishing behavior, even for low correlation. Finally, Fig. 8(b) demonstrates that SRA is the best precoding method under the evaluated system setup.

VII. Conclusion

This paper focused on the extension of the concept of space modulation to a VMIMO architecture. Based on the ideas of R-SM and SM, we formed an architecture which transfers information using the spatial position of multiple RNs, in addition, or not, to the conventional way. Using MIMO linear precoding, we were able to reduce or practically avoid the native problem of the activation of multiple RNs of the space modulated VMIMO schemes.

We compared our architecture against the appropriate SotA schemes and demonstrated significant BER gains. Especially, for a very large system setup and BER $= 10^{-4}$, the performance difference was up to 9 dB for DH-wBRS and up to 10 dB for Dh-w1R. Hence, we concluded that the concept of space modulation can be successfully applied in a distributed framework. Moreover, we provided strict bounds for the ABEP, both for the CD and DD algorithms at the RNs.

In order to provide linear precoding using a practical scenario of I-CSIT, we proposed two precoding methods. The first precoder is based on a statistical approach, while, the second precoder follows a worst case method. The instantaneous and average power of each precoder, used in this paper, was analyzed using a theoretical framework. The theoretical and simulated curves perfectly match. Using simulation results, we came to the conclusion that our architecture and precoding methods are robust to realistic scenarios of CSIT and correlation.

References

[1] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov./Dec. 1999.
[2] E. Larsson, "MIMO detection methods: How they work," *IEEE Signal Process. Mag.*, vol. 26, no. 3, pp. 91–95, May 2009.
[3] A. Stavridis, S. Sinanović, M. D. Renzo, and H. Haas, "Energy evaluation of spatial modulation at a multi-antenna base station," in *Proc. 78th IEEE VTC*, Las Vegas, NV, USA, Sep. 2–5, 2013, pp. 1–5.
[4] A. del Coso, U. Spagnolini, and C. Ibars, "Cooperative distributed MIMO channels in wireless sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 402–414, Feb. 2007.
[5] S. Jayaweera, "V-BLAST-based virtual MIMO for distributed wireless sensor networks," *IEEE Trans. Commun.*, vol. 55, no. 10, pp. 1867–1872, Oct. 2007.
[6] H. Chen, A. Gershman, and S. Shahbazpanahi, "Filter-and-forward distributed beamforming in relay networks with frequency selective fading," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1251–1262, Mar. 2010.
[7] L. Lai, K. Liu, and H. El-Gamal, "The three-node wireless network: Achievable rates and cooperation strategies," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 805–828, Mar. 2006.
[8] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037–3063, Sep. 2005.
[9] X. Hong, C.-X. Wang, M. Uysal, X. Ge, and S. Ouyang, "Capacity of hybrid cognitive radio networks with distributed VAAs," *IEEE Trans. Veh. Technol.*, vol. 59, no. 7, pp. 3510–3523, Sep. 2010.
[10] R. Mesleh, H. Haas, S. Sinanović, C. W. Ahn, and S. Yun, "Spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2228–2241, Jul. 2008.
[11] J. Jeganathan, A. Ghrayeb, and L. Szczecinski, "Spatial modulation: Optimal detection and performance analysis," *IEEE Commun. Lett.*, vol. 12, no. 8, pp. 545–547, Aug. 2008.
[12] M. Di Renzo, H. Haas, and P. M. Grant, "Spatial modulation for multiple-antenna wireless systems: A survey," *IEEE Commun. Mag.*, vol. 49, no. 12, pp. 182–191, Dec. 2011.
[13] J. Jeganathan, A. Ghrayeb, L. Szczecinski, and A. Ceron, "Space shift keying modulation for MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 8, no. 7, pp. 3692–3703, Jul. 2009.
[14] N. Serafimovski *et al.*, "Practical implementation of spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 62, no. 9, pp. 4511–4523, Nov. 2013.
[15] A. Younis *et al.*, "Performance of spatial modulation using measured real-world channels," in *Proc. 78th IEEE VTC*, Las Vegas, NV, USA, Sep. 2–5, 2013, pp. 1–5.
[16] J. Zhang, Y. Wang, L. Ding, and N. Zhang, "Bit error probability of spatial modulation over measured indoor channels," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1380–1387, Mar. 2014.
[17] M. Di Renzo, H. Haas, A. Ghrayeb, S. Sugiura, and L. Hanzo, "Spatial modulation for generalized MIMO: Challenges, opportunities, implementation," *Proc. IEEE*, vol. 102, no. 1, pp. 56–103, Jan. 2014.

[18] L.-L. Yang, "Transmitter preprocessing aided spatial modulation for multiple-input–multiple-output systems," in *Proc. 73rd IEEE VTC Spring*, May 2011, pp. 1–5.

[19] A. Stavridis, S. Sinanović, M. Di Renzo, and H. Haas, "Transmit precoding for receive spatial modulation using imperfect channel knowledge," in *Proc. 75th IEEE VTC Spring*, 2012, pp. 1–5.

[20] R. Mesleh, S. Ikki, and M. Alwakeel, "Performance analysis of space shift keying with amplify and forward relaying," *IEEE Commun. Lett.*, vol. 15, no. 12, pp. 1350–1352, Dec. 2011.

[21] M. Raed, I. Salama, A. Hadi, and A. Mansour, "Performance analysis of Space Shift Keying (SSK) modulation with multiple cooperative relays," *EURASIP J. Adv. Signal Process.*, vol. 2012, no. 1, p. 201, Sep. 2012.

[22] M. Wen, X. Cheng, V. Poor, and B. Jiao, "Use of Space Shift Keying (SSK) modulation in two-way amplify-and-forward relaying," *IEEE Trans. Veh. Technol.*, vol. 63, no. 3, pp. 1498–1504, Mar. 2014.

[23] P. Som and A. Chockalingam, "End-to-end BER analysis of space shift keying in decode-and-forward cooperative relaying," in *Proc. IEEE WCNC*, 2013, pp. 3465–3470.

[24] P. Som and A. Chockalingam, "BER analysis of space shift keying in cooperative multi-hop multi-branch DF relaying," in *Proc. 78th IEEE VTC*, Las Vegas, NV, USA, Sep. 2–5, 2013, pp. 1–5.

[25] S. Sugiura, S. Chen, and L. Hanzo, "Generalized space–time shift keying designed for flexible diversity-, multiplexing- and complexity-tradeoffs," *IEEE Trans. Wireless Commun.*, vol. 10, no. 4, pp. 1144–1153, Apr. 2011.

[26] S. Sugiura, S. Chen, H. Haas, P. Grant, and L. Hanzo, "Coherent versus non-coherent decode-and-forward relaying aided cooperative space–time shift keying," *IEEE Trans. Commun.*, vol. 59, no. 6, pp. 1707–1719, Jun. 2011.

[27] N. Serafimovski, S. Sinanovic, M. Di Renzo, and H. Haas, "Dual-hop Spatial Modulation (Dh-SM)," in *Proc. IEEE VTC Spring*, Budapest, Hungary, May 15–18, 2011, pp. 1–5.

[28] S. Narayanan, M. Di Renzo, F. Graziosi, and H. Haas, "Distributed spatial modulation for relay networks," in *Proc. 78th IEEE VTC*, Las Vegas, NV, USA, Sep. 2–5, 2013, pp. 1–6.

[29] S. Narayanan, M. Di Renzo, F. Graziosi, and H. Haas, "Distributed space shift keying for the uplink of relay-aided cellular networks," in *Proc. IEEE 17th Int. Workshop CAMAD Commun. Links Netw.*, 2012, pp. 130–134.

[30] Y. Yang and S. Aissa, "Information-guided transmission in decode-and-forward relaying systems: Spatial exploitation and throughput enhancement," *IEEE Trans. Wireless Commun.*, vol. 10, no. 7, pp. 2341–2351, Jul. 2011.

[31] P. Yang *et al.*, "Detect-and-forward relaying aided cooperative spatial modulation for wireless networks," *IEEE Trans. Commun.*, vol. 61, no. 11, pp. 4500–4511, Nov. 2013.

[32] S. Narayanan, A. Stavridis, M. Di Renzo, F. Graziosi, and H. Haas, "Distributed spatially-modulated space–time-block-codes," in *Proc. IEEE 18th Int. Workshop CAMAD Commun. Links Netw.*, 2013, pp. 159–163.

[33] A. Stavridis, S. Sinanović, M. D. Renzo, and H. Haas, "A power saving dual-hop architecture based on hybrid spatial modulation," in *Conf. Rec. 46th ASILOMAR Conf. Signals, Syst. Comput.*, Nov. 4–7, 2012, pp. 1366–1370.

[34] J. G. Proakis and M. Salehi, *Communication System Engineering*. Upper Saddle River, NJ, USA: Prentice-Hall, 1994.

[35] A. Wiesel, Y. Eldar, and S. Shamai, "Zero-forcing precoding and generalized inverses," *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4409–4418, Sep. 2008.

[36] C. Peel, B. Hochwald, and A. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication–Part I: Channel inversion and regularization," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 195–202, Jan. 2005.

[37] W. Ho, T. Quek, S. Sun, and R. Heath, "Decentralized precoding for multicell MIMO downlink," *IEEE Trans. Wireless Commun.*, vol. 10, no. 6, pp. 1798–1809, Jun. 2011.

[38] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ., 2004.

[39] A. Tulino, A. Lozano, and S. Verdu, "Impact of antenna correlation on the capacity of multiantenna channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2491–2509, Jul. 2005.

[40] A. M. Tulino and S. Verdu, "Random matrix theory and wireless communications," *Found. Trends Commun. Inf. Theory*, vol. 1, no. 1, pp. 1–182, Jun. 2004.

[41] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, Products*, A. Jeffrey and D. Zwillinger, Eds., 7th ed. San Diego, CA, USA: Academic, Mar. 2007.

[42] G. Alfano, A. Tulino, A. Lozano, and S. Verdu, "Capacity of MIMO channels with one-sided correlation," in *Proc. IEEE 8th Int. Symp. Spread Spectr. Tech. Appl.*, Aug. 2004, pp. 515–519.

[43] A. Mehana and A. Nosratinia, "Diversity of MIMO linear precoding," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 1019–1038, Feb. 2014.

[44] N. Jindal and A. Goldsmith, "Dirty-paper coding versus TDMA for MIMO broadcast channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 5, pp. 1783–1794, May 2005.

[45] M. O. Hasna and M.-S. Alouini, "End-to-end performance of transmission systems with relays over Rayleigh-fading channels," *IEEE Trans. Wireless Commun.*, vol. 2, no. 6, pp. 1126–1131, Nov. 2003.

[46] A. Younis, D. A. Basnayaka, and H. Haas, "Performance analysis for generalised spatial modulation," in *Proc. EW Conf.*, Barcelona, Spain, May 14–16, 2014, pp. 207–212.

[47] M. Di Renzo and H. Haas, "Bit error probability of space modulation over Nakagami-*m* fading: Asymptotic analysis," *IEEE Commun. Lett.*, vol. 15, no. 10, pp. 1026–1028, Oct. 2011.

[48] D. Gore, R. Heath, and A. Paulraj, "Transmit selection in spatial multiplexing systems," *IEEE Commun. Lett.*, vol. 6, no. 11, pp. 491–493, Nov. 2002.

[49] N. Balakrishnan and C. D. Lai, *Continuous Bivariate Distributions*. New York, NY, USA: Springer-Verlag, 2009.

[50] N. Serafimovski, S. Sinanovic, M. Renzo, and H. Haas, "Multiple access spatial modulation," *EURASIP J. Wireless Commun. Netw.*, vol. 2012, no. 1, pp. 299-1–299-20, Sep. 2012.

[51] F. Rusek *et al.*, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

Athanasios Stavridis (S'12) was born in Kastoria, Greece. He holds the Diploma degree (M.Eng. equivalent) in computer engineering and informatics and the M.Sc. degree in signal processing and communication systems from the University of Patras, Patras, Greece, in 2006 and 2008, respectively. He is currently working toward the Ph.D. degree with The University of Edinburgh, Edinburgh, U.K. He was awarded with a scholarship from the University of Patras upon the completion of his M.Sc. studies.

During the period of June 2011–June 2014, he was a Marie Curie Early-Stage-Researcher at The University of Edinburgh, Edinburgh, U.K. In January 2014, he was selected as an Exemplary Reviewer of the IEEE Communication Letters Reviewer Appreciation Program 2013. His main research interests include energy efficient physical layer techniques for wireless communications, spatial modulation, MIMO systems, and cooperative communications.

Dushyantha Basnayaka (S'11–M'12) received the B.Sc.Eng. degree (with first class honors) from the University of Peradeniya, Peradeniya, Sri Lanka, in 2006 and the Ph.D. degree in electrical engineering from the University of Canterbury, Christchurch, New Zealand, in 2012.

He was a System Engineer at MillenniumIT (a member company of the London Stock Exchange Group) from 2006 to 2009. He was with the Communication Research Group, University of Canterbury, from 2009 until 2012. He is now a Research Fellow in Wireless Communication at The University of Edinburgh, Edinburgh, U.K. He is the holder of one U.S. patent. His current research interests include massive MIMO, spatial modulation, interference mitigation techniques for cellular wireless systems, and macrodiversity multiuser MIMO systems.

Dr. Basnayaka was a recipient of University of Canterbury International Doctoral Scholarship for his doctoral studies at the University of Canterbury.

Sinan Sinanovic (S'98–M'07) received the Ph.D. degree in electrical and computer engineering from Rice University, Houston, TX, USA, in 2006. In 2006, he joined Jacobs University Bremen, Germany, as a Postdoctoral Fellow. In 2007, he joined The University of Edinburgh, U.K., where he was a Research Fellow at the Institute for Digital Communications. While working with Halliburton Energy Services, he has developed acoustic telemetry receiver, which was patented. He has also worked for Texas Instruments on development of ASIC testing. He is currently a Lecturer at Glasgow Caledonian University, Glasgow, U.K. Dr. Sinanovic is a member of the Tau Beta Pi Engineering Honor Society and a member of Eta Kappa Nu Electrical Engineering Honor Society. He won an honorable mention at the International Math Olympiad in 1994.

Marco Di Renzo (S'05–A'07–M'09–SM'14) received the Laurea (*cum laude*) and Ph.D. degrees in electrical and information engineering from the University of L'Aquila, L'Aquila, Italy, in 2003 and 2007, respectively, and the Habilitation à Diriger des Recherches from the University of Paris-Sud XI, Paris, France, in 2013.

Since January 2010, he has been a Tenured Academic Researcher ("Chargé de Recherche Titulaire") with the French National Center for Scientific Research (CNRS), as well as a Faculty Member of the Laboratory of Signals and Systems (L2S), a joint research laboratory of the CNRS, the École Supérieure d'Électricité (SUPÉLEC), and the University of Paris-Sud XI. His main research interests are in the area of wireless communications theory.

Dr. Di Renzo currently serves as an Editor of the IEEE COMMUNICATIONS LETTERS and the IEEE TRANSACTIONS ON COMMUNICATIONS (Heterogeneous Networks Modeling and Analysis). He was a recipient of several awards, which include a special mention for the outstanding five-year (1997–2003) academic career, University of L'Aquila; the THALES Communications Fellowship (2003–2006), University of L'Aquila; the 2004 Best Spin-Off Company Award, Abruzzo Region, Italy; the 2006 DEWS Travel Grant Award, University of L'Aquila; the 2008 Torres Quevedo Award, Ministry of Science and Innovation, Spain; the "Dérogation pour l'Encadrement de Thèse" (2010), University of Paris-Sud XI; the 2012 IEEE CAMAD Best Paper Award; the 2012 IEEE WIRELESS COMMUNICATIONS LETTERS Exemplary Reviewer Award; the 2013 IEEE VTC-Fall Best Student Paper Award; the 2013 Network of Excellence NEWCOM# Best Paper Award; the 2013 IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY Top Reviewer Award; the 2013 IEEE-COMSOC Best Young Researcher Award for Europe, Middle East and Africa (EMEA Region); and the 2014 IEEE ICNC Single Best Paper Award Nomination (Wireless Communications Symposium).

Harald Haas received the Ph.D. degree from The University of Edinburgh, Edinburgh, U.K., in 2001. He holds the Chair of Mobile Communications at The University of Edinburgh, Edinburgh, U.K. He has published 280 conference and journal papers, including a paper in *Science*. He holds 26 patents and has more than 20 pending patent applications. His main research interests are in optical wireless communications, hybrid optical wireless and RF communications, spatial modulation and massive MIMO, and interference coordination in wireless networks. Prof. Haas was an invited speaker at TED Global 2011, and his talk has been watched online more than 1.5 million times. He is a Co-Founder and a Chief Scientific Officer of pureLiFi Ltd. He first introduced and coined "spatial modulation" and "Li-Fi". Li-Fi was listed among the 50 best inventions in *TIME Magazine* 2011. He was co-recipient of a Best Paper Award at the IEEE Vehicular Technology Conference in Las Vegas in 2013. In 2012, he was the only recipient of the prestigious Established Career Fellowship from the Engineering and Physical Sciences Research Council (EPSRC) within Information and Communications Technology in the U.K. He was a recipient of the Tam Dalyell Prize 2013 awarded by The University of Edinburgh for excellence in engaging the public with science. In 2014, he was selected by EPSRC as one of ten RISE (Recognising Inspirational Scientists and Engineers) Leaders.

F.2 IEEE Transactions on Wireless Communications

Performance Analysis of Multi-Stream Receive Spatial Modulation in the MIMO Broadcast Channel

Athanasios Stavridis, Marco Di Renzo, *Senior Member, IEEE*, and Harald Haas, *Member, IEEE*

***Abstract*—In this paper, Multi-Stream Receive-Spatial Modulation (MSR-SM) for application to the Multiple-Input Multiple-Output (MIMO) broadcast channel is introduced and studied. MSR-SM is a closed-loop transmission scheme, which applies the concept of multi-stream space modulation at the receiver side. An accurate mathematical framework for the evaluation of the Bit Error Rate (BER) is proposed. In addition, the diversity order and coding gain of the new architecture are derived. Note that the proposed analytical framework takes into account both the small-scale fading and the system topology, and is directly applicable to the conventional MIMO broadcast channel. Compared with the state-of-the-art MIMO transmission in the broadcast channel, it is mathematically shown that MSR-SM achieves the same diversity order and a better coding gain, in the high Signal-to-Noise Ratio (SNR) regime. Finally, the proposed mathematical framework and the new findings are validated via Monte Carlo simulation results.**

I. Introduction

MULTI-ANTENNA communication has been considered as a promising technique for achieving high data rates without requiring additional radio resources [1, 2]. However, due to the deployment of multiple antennas, the complexity of the transceiver could become prohibitively high. A Multiple-Input Multiple-Output (MIMO) scheme which promotes a low complexity implementation is Spatial Modulation (SM) [3–7]. Due to its operating mechanism, SM requires a single Radio Frequency (RF) front-end at the transmitter [3]. This is shown to offer significant energy gains compared with conventional MIMO techniques [8]. In addition, at the receiver side, a low complexity (single stream) Maximum Likelihood (ML) detector is deployed [7]. Despite the adoption of a single stream detector, SM is able to obtain a multiplexing gain.

Inspired by the potential of SM, several authors have extended the concept of SM in different communication scenarios [9–19]. For example, Space Shift Keying (SSK) is a low complexity and low rate variant of SM [9]. Furthermore, Space Time Shift Keying (STSK) is a SM based scheme which extends the concept of SM in the time domain [20]. The first real system implementation of SM has recently been reported in [21]. The performance of SM under real channel measurements is discussed in [22, 23]. A complete introduction on SM is provided in [3].

Parts of this work will be presented at the 2015 IEEE Global Communications Conference (GLOBECOM), San Diego, California, USA, December 2015.

A. Stavridis and H. Haas are with the Li-Fi Research and Development Centre, Institute for Digital Communications, School of Engineering, The University of Edinburgh, Edinburgh, UK. E-mail: {a.stavridis, h.haas}@ed.ac.uk.

M. Di Renzo is with Paris-Saclay University, Laboratory of Signals and Systems (UMR-8506), CNRS - CentraleSupelec - University Paris-Sud XI, 91192 Gif-sur-Yvette (Paris), France (e-mail: marco.direnzo@lss.supelec.fr).

A. Related Work and Motivation

Similar to the concept of conventional SM, the author of [24] proposes a point-to-point closed-loop MIMO scheme that applies the principle of SM at the receiver side. In particular, using MIMO linear precoding, the reciprocal of SM, called Receive-Spatial Modulation (R-SM), is obtained. The extension of R-SM to a scheme which spatially modulates multiple parallel symbol streams to the indices of multiple receiving antennas is conducted in [25, 26]. In the present paper, the term of Multi-Stream Receive-Spatial Modulation (MSR-SM) is used for this scheme. In addition, the performance of R-SM in different application scenarios is studied in [27–29].

As discussed in [3], there is a wide range of spatially modulated architectures for point-to-point communication. However, *is it possible for SM to be incorporated in a Multi-User (MU) scenario?* Indeed, just like any other MIMO physical layer technique, SM and its variants can be combined with a multiple access scheme such as Time Division Multiple Access (TDMA), Frequency Division Multiple Access (FDMA), or Orthogonal Frequency-Division Multiple Access (OFDMA) in order to form a MU system.

A new trend in MIMO communication promotes systems where multiple users are aggressively allocated in the same time and frequency resources. Usually, this is accomplished via Space Division Multiple Access (SDMA) techniques [30, 31]. However, the design of a SDMA-based spatially modulated scheme, is a challenging task. Due to the activation of a single transmit antenna (or a subset of the available transmit antennas) and the way that information is conveyed, the design of interference reduction, elimination, or manipulation techniques becomes difficult. However, the authors of [32, 33] managed to incorporate SM in the MIMO broadcast channel. In these papers, non-interfering SM-based communication is established via the use of a linear precoding matrix which is based on the Zero Forcing (ZF) principle.

B. Contributions and Outcomes

Against this background, the present paper aims to incorporate MSR-SM in the MIMO broadcast channel. More specifically, a new SDMA-based architecture based on the concept of MSR-SM is proposed. In addition, an accurate mathematical framework for computing the Average Bit Error Probability (ABEP), the diversity order, and the coding gain is introduced. The proposed transmission scheme and framework, in particular, are based on the following assumptions: i) the wireless channel follows a Rayleigh distribution; ii) Perfect-Channel State Information at the Transmitter (P-CSIT) is assumed; iii) the proposed precoder is designed based on

the ZF principle; and iv) the effect of system topology is duly taken into account. Based on the proposed mathematical framework, it is proved that the proposed scheme is capable of outperforming the conventional MIMO broadcast channel. In particular, the proposed scheme provides the same diversity order as state-of-the-art MIMO schemes but a better coding gain in the high Signal-to-Noise Ratio (SNR) regime.

As far as the novelty of the proposed mathematical framework is concerned, to the best of the authors' knowledge, the analysis of the diversity order and of the coding gain of MSR-SM is not available in the literature. In the present paper, we prove, for the first time, that that the conventional MIMO broadcast channel and MU MSR-SM achieve the same diversity order if they both rely on ZF precoding. In addition, the proposed approach is directly applicable to point-to-point single user scenarios. As far as this latter scenario is concerned, in particular, it is worth mentioning that the ABEP of MSR-SM for point-to-point single user transmission has recently been studied in [25, 26]. However, several important differences exist between the framework available in [25, 26] and that proposed in the present paper. The study presented in [25] does not take into account the statistical description of the received signal. The ABEP computed in [26] is applicable to MSR-SM in the presence of a suboptimal detector, which decouples the detection process. In contrast, the analysis presented in Sections IV and V is different for the following reasons: i) the statistical description of the received signal is considered; ii) the system topology that is inherent in MU setups is taken into account; and iii) the detection process is based on the ML principle which imposes some additional mathematical difficulties. Finally, since the conventional Spatial MultipleXing (SMX) MIMO architecture with ZF precoding is a special case of MSR-SM, the proposed framework can be applied to this setup as well and can be used for a simple comparison between the two architectures, as better discussed in the sequel.

The rest of the present paper is organized as follows. In Section II, the system model is introduced. In Section IV, the ABEP of each user and of the whole system is computed. In Section V, the diversity order and coding gain of MSR-SM in the MIMO broadcast channel are analyzed. In Section VI, the proposed MSR-SM MIMO architecture is compared against the corresponding conventional MIMO broadcast channel and some numerical results that validate our theoretical findings are illustrated. Finally, Section VII concludes this paper.

Notation: Lowercase bold letters denote vectors and uppercase bolt letters denote matrices. $(\cdot)^T$, $(\cdot)^H$, tr$(\cdot)$ and $\mathbf{A}^{1/2}$ denote transpose, Hermitian transpose, matrix trace and the square root of $\mathbf{A}$, respectively. The Kronecker product is denoted as $\otimes$. $\|\cdot\|_2$ represents the Euclidean norm, while $\|\cdot\|_F$ is the Frobenius norm. A diagonal matrix, whose main diagonal includes the elements $a_1, \cdots, a_n$, is denoted as $\mathbf{diag}\left(a_1, \ldots, a_n\right)$. E$[\cdot]$ denotes the mean value of a RV. A complex Gaussian distribution with mean m and variance σ_C^2 is represented as $\mathcal{CN}(m, \sigma_C^2)$, where its real and imaginary part are independent and identically distributed (i.i.d.) Gaussian RV with distribution $\mathcal{N}(m, \frac{\sigma_C^2}{2})$. Re$\{\cdot\}$ denotes the real part of a complex number or matrix.

II. System Model

An uncoded Multi-User Multiple-Input Multiple-Output (MU-MIMO) system that comprises a multi-antenna Base Station (BS) and N_u remotely distributed multi-antenna users is considered. The BS is equipped with N_t antennas and each user possesses N_r antennas. Since the transmitter is a BS, it is realistic to consider the assumption that $N_t \geq N_u N_r$. In addition, the wireless channel between the BS and every user is assumed to be frequency flat and quasi-static. Finally, P-CSIT is considered, which can be obtained by using either the channel reciprocity or fast and error free links from the users[1].

Provided the availability of P-CSIT, the transmitter is able to use linear precoding. By interpreting the N_t transmit antennas and the $B = N_u N_r$ receive antennas as a $N_t \times B$ MIMO system, the baseband equation of such a system is expressed as:

$$\mathbf{y} = \breve{\mathbf{H}}\breve{\mathbf{P}}\breve{\mathbf{D}}\mathbf{x} + \mathbf{w}, \tag{1}$$

in a matrix form. In (1), $\mathbf{y} = \left[\mathbf{y}_1^T, \ldots, \mathbf{y}_{N_u}^T\right]^T$ is a $N_u N_r \times 1$ vector, where $\mathbf{y}_i$, $i = 1, \ldots, N_u$, denotes the $N_r \times 1$ received signal vector at the i-th user. The $N_u N_r \times N_t$ matrix, $\breve{\mathbf{H}} = \left[\breve{\mathbf{H}}_1^H, \ldots, \breve{\mathbf{H}}_{N_u}^H\right]^H$, denotes the wireless channel from the transmitter to all receive antennas. Furthermore, the submatrix $\breve{\mathbf{H}}_i$, $i = 1, \ldots, N_u$, denotes the wireless channel from the transmitter to the i-th user. Due to the spatial distribution of the users inside the geographical area of a cell, each user experiences a different large-scale channel effect. However, given that the receive antennas of each user are collocated, the large-scale channel effect between the transmitter and each receive antenna of a certain user is the same. Therefore, it is assumed that $\breve{\mathbf{H}}_i$, $i = 1, \ldots, N_u$, has the following distribution: $\breve{\mathbf{H}}_i \sim \mathcal{CN}(\mathbf{0}, \xi_i \mathbf{I})$. Here, the value of ξ_i is determined by the transmission distance and the effect of shadowing. In general, the values of ξ_i close to zero represent a poor channel condition, while those of ξ_i close to one indicate a strong channel condition. It is assumed that there is no channel correlation due to rich scattering. The $N_t \times N_u N_r$ precoding matrix can be formulated as $\breve{\mathbf{P}} = \left[\breve{\mathbf{P}}_1, \ldots, \breve{\mathbf{P}}_{N_u}\right]$, where, $\breve{\mathbf{P}}_i$, $i = 1, \ldots, N_u$, corresponds to the precoding matrix of the i-th user. In order to ensure that the transmitted power is not amplified by the precoder $\breve{\mathbf{P}}$, a $N_u N_r \times N_u N_r$ diagonal normalization matrix $\breve{\mathbf{D}} = \mathbf{diag}\left(\breve{d}_1, \ldots, \breve{d}_{N_u N_r}\right)$ is used. Every element $\breve{d}_i$, $i = 1, \ldots, N_u N_r$, of $\breve{\mathbf{D}}$ is expressed as, $\breve{d}_i = \sqrt{1/\|\breve{\mathbf{p}}_i\|_2^2}$, where, $\breve{\mathbf{p}}_i$ is the i-th column of $\breve{\mathbf{P}}$. Thus, every column of the normalized precoding matrix, $\breve{\mathbf{P}}_{\text{norm}} = \breve{\mathbf{P}}\breve{\mathbf{D}}$ has unity power. The normalization matrix $\breve{\mathbf{D}}$ can be expressed in the following block diagonal matrix form, $\breve{\mathbf{D}} = \mathbf{diag}\left(\breve{\mathbf{D}}_1, \ldots, \breve{\mathbf{D}}_{N_u}\right)$. Here, $\breve{\mathbf{D}}_i$, $i = 1, \ldots, N_u$, is the $N_r \times N_r$ diagonal normalization matrix of the corresponding precoding matrix $\breve{\mathbf{P}}_i$. The collective signal vector at the

[1] In real systems, Channel State Information at the Transmitter (CSIT) is subjected to imperfections. However, the study of the effect of imperfect CSIT is out of the scope of this paper.

$$\mathbf{x}_i = \left[0,\ldots,0,\ \underbrace{s_1}_{i_1\text{-th position}},0,\ldots,0,\ \underbrace{s_i}_{i_k\text{-th position}},0,\ldots,0,\ \underbrace{s_{N_s}}_{i_{N_s}\text{-th position}}\ 0\ldots,0\right]^T \tag{7}$$

transmitter is denoted as, $\mathbf{x} = \left[\mathbf{x}_1^T,\ldots,\mathbf{x}_{N_u}^T\right]^T$, where, $\mathbf{x}_i$, $i = 1,\ldots,N_u$, is the signal vector for the i-the user. Finally, $\mathbf{w} = \left[\mathbf{w}_1^T,\ldots,\mathbf{w}_{N_u}^T\right]^T \sim \mathcal{CN}(\mathbf{0},\sigma_{\mathbf{w}}^2\mathbf{I})$ is a $N_uN_r \times 1$ vector that represents the white Gaussian noise. In more detail, $\mathbf{w}_i$, $i = 1,\ldots,N_u$, is the Gaussian noise observed by the i-th user.

The precoding method of interest in the present paper is ZF. The ZF precoder is a suboptimal precoder that offers a good trade-off between complexity and performance [34]. The main characteristic of ZF precoding is the total elimination of interference between different users and between different antennas of the same user. Hence, ZF precoding is an efficient method that can be used for the formation of a MU architecture based on MSR-SM. Therefore, if the channel matrix is expressed as,

$$\breve{\mathbf{H}} = \mathbf{\Xi}^{\frac{1}{2}}\mathbf{H}, \tag{2}$$

where, $\mathbf{\Xi} = \mathbf{diag}\left(\xi_1\mathbf{I},\ldots,\xi_{N_u}\mathbf{I}\right)$, represents the effect of the system topology, $\mathbf{H} = \left[\mathbf{H}_1^H,\ldots,\mathbf{H}_{N_u}^H\right]^H \sim \mathcal{CN}(\mathbf{0},\mathbf{I})$, represents the small-scale fading, and $\mathbf{H}_i$ represents the small scale effect of the i-th user, the ZF precoding matrix is written as:

$$\breve{\mathbf{P}} = \mathbf{H}^H\left(\mathbf{H}\mathbf{H}^H\right)^{-1}\mathbf{\Xi}^{-\frac{1}{2}}. \tag{3}$$

Let the ZF precoder in (3), the diagonal normalization matrix $\breve{\mathbf{D}}$ can be re-written as, $\breve{\mathbf{D}} = \mathbf{\Xi}^{\frac{1}{2}}\mathbf{D}_{\text{MU}}$, where, $\mathbf{D}_{\text{MU}} = \mathbf{diag}\left(d_1,\ldots,d_{N_uN_r}\right)$ is a diagonal matrix. The i-th element of the main diagonal of $\mathbf{D}_{\text{MU}}$ is expressed as:

$$d_i = \sqrt{\frac{1}{\left[\left(\mathbf{H}\mathbf{H}^H\right)^{-1}\right]_{i,i}}},\ i = 1,\ldots,N_uN_r, \tag{4}$$

in order to ensure that the instantaneous transmission power is constrained.

In order to gain a better understanding, (2) and (3) can be plugged into (1). By doing so, the received signal at user is expressed as:

$$\mathbf{y}_i = \sqrt{\xi_i}\mathbf{D}_i\mathbf{x}_i + \mathbf{w}_i,\quad i = 1,\ldots,N_u, \tag{5}$$

since $\mathbf{D}_{\text{MU}} = \mathbf{diag}\left(\mathbf{D}_1\ldots,\mathbf{D}_{N_u}\right)$ is a block diagonal matrix, where $\mathbf{D}_i$ is defined as the i-th $N_r \times N_r$ block matrix of the main block diagonal of $\mathbf{D}_{\text{MU}}$. From (5), it can be seen that the received signal of the i-th user is directly affected by the CSIT of all users.

The inspection of (5) highlights that the choice of the transmitted vector $\mathbf{x}_i$, $i = 1,\ldots,N_u$, determines the way that information is transmitted to each user. For example, conventional SMX transmission is obtained if all the elements of $\mathbf{x}_i$ are drawn from a conventional M-ary constellation diagram $\mathcal{M}$. On the other hand, MSR-SM can be obtained by appropriately choosing the symbol vectors $\mathbf{x}_i$, $i = 1,\ldots,N_u$, as discussed in Section IV.

At the users' side, the reconstruction of the transmitted bit-streams is undertaken by detecting the transmitted vectors $\mathbf{x}_i$, $i = 1,\ldots,N_u$. Provided that the i-th user is aware of ξ_i and $\mathbf{D}_i$, this can be implemented at every user independently by using the following ML detector:

$$(\tilde{\mathbf{x}}_i) = \arg\min_{\mathbf{x}_i}\|\mathbf{y}_i - \sqrt{\xi_i}\mathbf{D}_i\mathbf{x}_i\|_2^2,\quad i = 1,\ldots,N_u. \tag{6}$$

III. Multi-Stream Receive-Spatial Modulation

The objective of MSR-SM is twofold: i) the simultaneous transmission of $N_s \leq N_r$ information symbols from the transmitter to the receiver and ii) the transmission of additional information bits via the indices of N_s (out of N_r) receive antennas. Assuming ZF precoding, the received signal at each user is given in (5). Hence, by appropriately choosing the structure of the transmit signal vector $\mathbf{x}_i$, it is possible to enforce that the noise free received signal $\mathbf{D}_i\mathbf{x}_i$ has exactly N_s non-zero elements and $N_r - N_s$ zero elements. Let $\mathbf{D}_i$ be the diagonal matrix introduced in (5), the non-zero elements of $\mathbf{D}_i\mathbf{x}_i$ constitute a scaled version of the corresponding non-zero elements of $\mathbf{x}_i$. Similarly, the positions of the zero elements of $\mathbf{D}_i\mathbf{x}_i$ are the same as those of the zero elements of $\mathbf{x}_i$. This implies that a portion of binary information can be encoded in the position of the non-zero elements of $\mathbf{x}_i$ and consequently in the position of the non-zero elements of the noise free received signal $\mathbf{D}_i\mathbf{x}_i$.

The general expression of $\mathbf{x}_i \in \mathcal{B}_i$ for MSR-SM is given in (7), which is available at the top of this page. Here, $\mathcal{B}_i$ denotes the set (alphabet) of all possible transmitted symbol vectors to the i-th user and $\{s_1,\ldots,s_{N_s}\} \in \mathcal{M}$, where $\mathcal{M}$ is the deployed constellation. The positions of the non-zero elements correspond to the indices of the receiving antennas, while the positions of the zero elements correspond to the antennas that do not receive signal. The selection of the combinations of receiving antennas can be optimized in order to minimize the instantaneous Bit Error Rate (BER) or they can be chosen at random. In this paper, for reasons of simplicity and mathematical tractability, the focus is on the latter case.

With these assumptions, the bit-stream to be transmitted in every single period is divided in two portions. The first portion, which is of length $k_1^{\text{MSR-SM}} = N_s\log_2(M)$ bits, is encoded and transmitted using the N_s symbols which are drawn from the M-ary constellation $\mathcal{M}$. The second portion, which is of length $k_2^{\text{MSR-SM}} = \log_2\left(\lfloor\binom{N_r}{N_s}\rfloor\right)$ bits, is encoded in the indices of the receive antennas. Therefore, the spectral efficiency of MSR-SM is $k_{\text{MSR-SM}} = N_s\log_2(M) + \log_2\left(\lfloor\binom{N_r}{N_s}\rfloor\right)$ bits per channel use (bpcu) per user. In the extreme case, where it holds $N_s = N_r$, MSR-SM reduces to a spatially multiplexed MIMO architecture with ZF precoding. In this case, the spectral efficiency is $k_{\text{SMX}} = N_r\log_2(M)$ bpcu per user.

The inspection of (7) reveals that the sparsity of the transmission alphabet $\mathcal{B}_i$ of MSR-SM can be utilized in order to offer a low computational complexity at the transmitter. Provided that the precoding matrix $\breve{\mathbf{P}}$ of (3) is precomputed offline before transmission, the transmitted signal, $\mathbf{s} = \breve{\mathbf{P}}\breve{\mathbf{D}}\mathbf{x}$, in (1) can be computed with $\mathcal{C}_t = N_t(8N_uN_s - 2) + 2N_uN_s$ real operations (additions or multiplications). It is clear that as N_s takes lower values, the complexity of the transmitter $\mathcal{C}_t$ is also reduced. The computational analysis of the detector of MSR-SM is presented in [26]. In that paper, a suboptimal detector with low complexity for MSR-SM is also proposed. The study of the detector of [26] is, however, outside of the scope of the present paper.

IV. Evaluation of the Average Bit Error Probability

In this section, the ABEP of each individual user and of the whole MU system are derived. The ABEP of the i-th user, $P^i_{\text{bit}}(\gamma)$, for a given transmit SNR γ, can be bounded as follows:

$$P^i_{\text{bit}}(\gamma) \leq \frac{1}{|\mathcal{B}_i| k_{\text{MSR-SM}}} \sum_{\mathbf{x}_i} \sum_{\substack{\hat{\mathbf{x}}_i \\ \hat{\mathbf{x}}_i \neq \mathbf{x}_i}} d(\mathbf{x}_i \to \hat{\mathbf{x}}_i) P^i_{\text{e}}(\mathbf{x}_i \to \hat{\mathbf{x}}_i, \gamma), \tag{8}$$

using the union bound technique [35]. In (8), $P^i_{\text{e}}(\mathbf{x}_i \to \hat{\mathbf{x}}_i, \gamma)$ represents the Pairwise Error Probability (PEP) of transmitting $\mathbf{x}_i$ to the i-th user while the detector decides in favor of the erroneous symbol vector $\hat{\mathbf{x}}_i$. The number of different bits between the bit-words represented by $\mathbf{x}_i$ and $\hat{\mathbf{x}}_i$ is denoted by $d(\mathbf{x}_i \to \hat{\mathbf{x}}_i)$. Furthermore, $|\mathcal{B}_i| = M^{N_s} 2^{\left\lfloor \log_2 \binom{N_r}{N_s} \right\rfloor}$ denotes the number of all possible transmitted symbol vectors to the i-th user.

The evaluation of (8) requires the knowledge of $P^i_{\text{e}}(\mathbf{x}_i \to \hat{\mathbf{x}}_i, \gamma)$, which is the expectation of the instantaneous PEP over all channel realizations. Let the detector of the i-th user in (6), a symbol error occurs at this user when, $\mathcal{E}_i(\mathbf{x}_i, \hat{\mathbf{x}}_i) = \left\{ \|\mathbf{y}_i - \sqrt{\xi_i}\mathbf{D}_i\mathbf{x}_i\|_2^2 > \|\mathbf{y}_i - \sqrt{\xi_i}\mathbf{D}_i\hat{\mathbf{x}}_i\|_2^2 \right\}$. In this case, if the statistical distribution of the Gaussian noise of the i-th user is taken into account, after some manipulations, the corresponding instantaneous PEP (conditioned on $\mathbf{D}_i$) is expressed as:

$$P^i_{\text{e}}\left(\mathbf{x}_i \to \hat{\mathbf{x}}_i, \gamma | \mathbf{D}_i^2\right) = Q\left(\sqrt{\frac{\mathbf{c}_i^H \mathbf{D}_i^2 \mathbf{c}_i}{2} \xi_i \gamma}\right). \tag{9}$$

In (9), the vector $\mathbf{c}_i$ is defined as $\mathbf{c}_i = \mathbf{x}_i - \hat{\mathbf{x}}_i$. From (9) and the structure of $\mathbf{D}_i$, it can be seen that the instantaneous PEP of the i-th user depends on the CSIT of all users via $\mathbf{D}_i$.

For notational convenience, the following variables are defined:

$$z_i = \mathbf{c}_i^H \mathbf{D}_i^2 \mathbf{c}_i, \tag{10}$$

and

$$\breve{\gamma}_i = \xi_i \gamma. \tag{11}$$

In order to evaluate the expectation of (9) over all possible realizations of the diagonal random matrix $\mathbf{D}_i$, (10), (11), and the following tight upper bound of the Q-function [36], $Q(x) \leq \frac{1}{6}e^{-2x^2} + \frac{1}{12}e^{-x^2} + \frac{1}{4}e^{\frac{-x^2}{2}}$, are considered. In this way, the PEP of interest is expressed as:

$$P^i_{\text{e}}(\mathbf{x}_i \to \hat{\mathbf{x}}_i, \breve{\gamma}_i) \leq \frac{1}{6}\text{E}_{z_i}\left[e^{-z_i\breve{\gamma}_i}\right] + \frac{1}{12}\text{E}_{z_i}\left[e^{-\frac{z_i}{2}\breve{\gamma}_i}\right] + \frac{1}{4}\text{E}_{z_i}\left[e^{-\frac{z_i}{4}\breve{\gamma}_i}\right]. \tag{12}$$

From (12), it can be observed that the Probability Density Function (PDF) of the Random Variable (RV) z_i has to be derived. To this end, using an algebraic elaboration on (10), the RV z_i can be re-written as:

$$z_i = \sum_{k=1}^{N_r} |x_k - \hat{x}_k|^2 d_k^2 = \sum_{x_k - \hat{x}_k \neq 0} |x_k - \hat{x}_k|^2 d_k^2, \tag{13}$$

where, x_k and $\hat{x}_k$, $k = 1, \ldots, N_r$, are the k-th elements of $\mathbf{x}_i$ and $\hat{\mathbf{x}}_i$, $i = 1, \ldots, N_u$, respectively. Furthermore, d_k, $k = 1, \ldots, N_r$, is the k-th element of the main diagonal of $\mathbf{D}_i$.

Usually, in the literature, the RVs d_k^2 are assumed to be statistically independent in order to simplify the mathematical analysis [37, 38]. This assumption is, however, in contradiction with the structure of $d_k^2 = 1/\left[\left(\mathbf{H}_i\mathbf{H}_i^H\right)^{-1}\right]_{k,k}$. In fact, the realization of every RV d_k^2 occurs using the same mathematical operations on the same random matrix $\mathbf{H}_i$. More specifically, the following holds [39]:

$$d_k^2 = \frac{1}{\left[\text{adj}\left(\mathbf{H}_i\mathbf{H}_i^H\right)\right]_{k,k}} \det\left(\mathbf{H}_i\mathbf{H}_i^H\right), \tag{14}$$

where, adj$(\cdot)$ is the adjoint matrix and $\det(\cdot)$ is the matrix determinant. The inspection of (14) shows that for different values of $k = 1, \ldots, N_r$, the realization of the RVs d_k^2 affects one another, since they are produced via the same mathematical formula using the same random elements of $\mathbf{H}_i$. This implies that the RVs d_k^2 are dependent. An empirical confirmation for the previous argument can be obtained by computing the Pearson product-moment correlation coefficient between any pair of the previous RVs using multiple samples. In this way, it can be shown that the Pearson product-moment correlation coefficient takes non-zero values. The analytical evaluation of these correlation coefficients is difficult to be obtained, since it requires the joint PDF between each pair of the RVs d_k^2. An additional confirmation is provided in Section VI-A, where the empirical PDF of z_i is depicted against the theoretical PDF derived below. Hence, it can be concluded that the RVs d_k^2 are statistically dependent.

Due to the dependence between the RVs d_k^2, a different approach, compared to the state-of-the-art literature, is proposed: we take into account that d_k^2, $k = 1, \ldots, N_r$, are dependent and correlated gamma RVs.

For notational convenience, the variables

$$b_j = |c_k|^2 = |x_k - \hat{x}_k|^2, \tag{15}$$

and

$$X_j = d_k^2, \tag{16}$$

$j = 1, \ldots, N_i$, are introduced only for those values of k in (13) for which it holds that $c_k = x_k - \hat{x}_k \neq 0$. Here, c_k is

$$f_{z_i}(x) = \left[\prod_{l=1}^{N_i}\left(\frac{\breve{\alpha}_1}{\breve{\alpha}_l}\right)^{L_{\text{MU}}}\right]\left[\sum_{k=0}^{+\infty}\frac{\breve{\delta}_k x^{N_i L_{\text{MU}}+k-1}e^{-\frac{x}{\breve{\alpha}_1}}}{\breve{\alpha}_1^{N_i L_{\text{MU}}+k}\Gamma\left(N_i L_{\text{MU}}+k\right)}\right]H_0(x), \tag{21}$$

$$\breve{\delta}_{k+1} = \begin{cases} 1, & k=-1, \\ \frac{k}{k+1}\sum_{i=1}^{k+1}\left[\sum_{j=1}^{N}\left(1-\frac{\breve{\alpha}_1}{\breve{\alpha}_j}\right)^i\right]\breve{\delta}_{k+1-i}, & k=0,1,2,\ldots. \end{cases} \tag{25}$$

the k-th element of $\mathbf{c}_i$. In addition, N_i is the number of non zero elements of $\mathbf{c}_i$. Thus, the value of N_i depends on the considered pair of $\mathbf{x}_i$ and $\hat{\mathbf{x}}_i$. Therefore, (13) can be re-written as:

$$z_i = \sum_{j=1}^{N_i} b_j X_j = \sum_{j=1}^{N_i} Z_j. \tag{17}$$

In (17), Z_j is defined as:

$$Z_j = b_j X_j. \tag{18}$$

The PDF of d_k^2 is explicitly derived in [39] as a gamma distribution with $d_k^2 \sim \text{Gamma}\left(L_{\text{MU}}, 1\right)$ and

$$L_{\text{MU}} = N_t - N_u N_r + 1. \tag{19}$$

Consequently, given that $X_j = d_k^2$, X_j follows the same distribution. Therefore, the RVs $Z_j = b_j X_j$, $j = 1, \ldots, N_j$, are distributed as $Z_j \sim \text{Gamma}\left(L_{\text{MU}}, b_j\right)$ with a PDF given by:

$$f_{Z_j}(x) = \frac{1}{b_j^{L_{\text{MU}}}\Gamma(L_{\text{MU}})}x^{L_{\text{MU}}-1}e^{\frac{x}{b_j}}H_0(x), \tag{20}$$

Here, $H_0(x)$ is the Heaviside step function defined as $H_0(x) = 0$ for $x < 0$ and $H_0(x) = 1$ for $x \geq 0$.

Since X_j, $j = 1, \ldots, N_j$, are correlated RVs, also, $Z_j = b_j X_j$, $j = 1, \ldots, N_j$ are correlated RVs. This implies that z_i is a RV which is equal to the sum of correlated Gamma RVs. For this reason and based on [40, Corollary 1], the PDF of z_i is given in (21) at the top of this page. Note that, because of the correlation between the different pairs of RVs Z_j, the derivations provided below are significantly different than the corresponding derivations of a conventional MIMO system which deploys ZF detection. Such an example is the performance analysis presented in [31]. In fact, in a conventional MIMO system with ZF detection, the detection process of the parallel symbol streams decouples and each parallel symbol stream can be detected independently. Therefore, the correlation between the RVs which represent the instantaneous receive SNR of each parallel symbol stream does not need to be considered.

In (21), $\breve{\alpha}_l$, $l = 1, \ldots, N_i$, are the eigenvalues of:

$$\breve{\mathbf{A}} = \breve{\mathbf{B}}\breve{\mathbf{R}} \tag{22}$$

in ascending order, and $\breve{\mathbf{B}}$ is the diagonal matrix

$$\breve{\mathbf{B}} = \mathbf{diag}\left(b_1, \ldots, b_{N_i}\right), \tag{23}$$

where b_l, $l = 1, \ldots, N_i$, is the square of the absolute value of the l-th non zero element of $\mathbf{c}_i$. Moreover, $\breve{\mathbf{R}}$ is a $N_i \times N_i$ matrix defined as:

$$\breve{\mathbf{R}} = \begin{bmatrix} 1 & \sqrt{\rho_c} & \cdots & \sqrt{\rho_c} \\ \sqrt{\rho_c} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \sqrt{\rho_c} \\ \sqrt{\rho_c} & \cdots & \sqrt{\rho_c} & 1 \end{bmatrix}, \tag{24}$$

where, ρ_c is the Pearson product-moment correlation coefficient between any pair of two different RVs of the main diagonal of $\mathbf{D}_i^2$. The inspection of the structure of d_k^2 shows that the Pearson product-moment correlation coefficient between every pair of two different RVs of the main diagonal of $\mathbf{D}_i^2$ takes the value of ρ_c. Finally, $\breve{\delta}_k$, $k = 0, 1, 2, \ldots$, are given in (25) at the top of this page.

Let the PDF of z_i given in (21), the evaluation of (12) over all possible realizations of z_i can be performed by evaluating expectations of the following form, $g\left(y\right) = \text{E}_{z_i}\left[e^{-y\breve{\gamma}_i z_i}\right]$. Here, y is a deterministic scalar. The evaluation of the previous expectation is given as:

$$\begin{aligned} g\left(y\right) &= \text{E}_{z_i}\left[e^{-y\breve{\gamma}_i z_i}\right] = \int_{-\infty}^{+\infty} e^{-y\breve{\gamma}_i x} f_{z_i}(x)dx \\ &= \left[\prod_{l=1}^{N_i}\left(\frac{\breve{\alpha}_1}{\breve{\alpha}_l}\right)^{L_{\text{MU}}}\right]\left[\int_0^{+\infty} e^{-y\breve{\gamma}_i x}\right. \\ &\times \left.\sum_{k=0}^{+\infty}\frac{\breve{\delta}_k x^{N_i L_{\text{MU}}+k-1}e^{-\frac{x}{\breve{\alpha}_1}}}{\breve{\alpha}_1^{N_i L_{\text{MU}}+k}\Gamma(N_i L_{\text{MU}}+k)}dx\right] \\ &= \left[\prod_{l=1}^{N_i}\left(\frac{\breve{\alpha}_1}{\breve{\alpha}_l}\right)^{L_{\text{MU}}}\right]\sum_{k=0}^{+\infty}\left[\frac{\breve{\delta}_k}{\breve{\alpha}_1^{N_i L_{\text{MU}}+k}\Gamma(N_i L_{\text{MU}}+k)}\right. \\ &\times \left.\int_0^{+\infty} x^{N_i L_{\text{MU}}1+k-1}e^{-\left(y\breve{\gamma}_i+\frac{1}{\breve{\alpha}_1}\right)x}dx\right], \end{aligned} \tag{26}$$

From the integration formula in [41, p.346, 3.381, 4], we have:

$$\int_0^{+\infty} x^{\nu-1}e^{-\mu x}dx = \mu^{-\nu}\Gamma(\nu), \tag{27}$$

where $\nu > 0$ and $\text{Re}\{\mu\} > 0$, and $\Gamma(\cdot)$ denotes the incomplete gamma function defined in [41, p. 899]. By plugging (27) in (26) and with the aid of some manipulations, we have:

$$\begin{aligned} g\left(y\right) &= \left[\prod_{l=1}^{N_i}\left(\frac{\breve{\alpha}_1}{\breve{\alpha}_l}\right)^{L_{\text{MU}}}\right]\left(y\breve{\alpha}_1\breve{\gamma}_i+1\right)^{-N_i L_{\text{MU}}} \\ &\times \sum_{k=0}^{+\infty}\breve{\delta}_k\left(y\breve{\alpha}_1\breve{\gamma}_i+1\right)^{-k}, \end{aligned} \tag{28}$$

Since $\breve{\gamma}_i = \xi_i\gamma$ and using the result from (28), the PEP of

$$P_{\mathrm{e}}^{i}(\mathbf{x}\to\hat{\mathbf{x}},\gamma)\leq\frac{\left[\prod_{l=1}^{N_i}\left(\frac{\breve{\alpha}_1}{\breve{\alpha}_l}\right)^{L_{\mathrm{MU}}}\right]}{6}\left(\breve{\alpha}_1\xi_i\gamma+1\right)^{-N_iL_{\mathrm{MU}}}\sum_{k=0}^{+\infty}\breve{\delta}_k\left(\breve{\alpha}_1\xi_i\gamma+1\right)^{-k}$$
$$+\frac{\left[\prod_{l=1}^{N_i}\left(\frac{\breve{\alpha}_1}{\breve{\alpha}_l}\right)^{L_{\mathrm{MU}}}\right]}{12}\left(\frac{\breve{\alpha}_1}{2}\xi_i\gamma+1\right)^{-N_iL_{\mathrm{MU}}}\sum_{k=0}^{+\infty}\breve{\delta}_k\left(\frac{\breve{\alpha}_1}{2}\xi_i\gamma+1\right)^{-k}$$
$$+\frac{\left[\prod_{l=1}^{N_i}\left(\frac{\breve{\alpha}_1}{\breve{\alpha}_l}\right)^{L_{\mathrm{MU}}}\right]}{4}\left(\frac{\breve{\alpha}_1}{4}\xi_i\gamma+1\right)^{-N_iL_{\mathrm{MU}}}\sum_{k=0}^{+\infty}\breve{\delta}_k\left(\frac{\breve{\alpha}_1}{4}\xi_i\gamma+1\right)^{-k}. \tag{29}$$

$$P_i^{+\infty}(\mathbf{x}_i\to\hat{\mathbf{x}}_i,\gamma)\lessapprox\frac{\left[\prod_{l=1}^{N_i}\left(\frac{\breve{\alpha}_1}{\breve{\alpha}_l}\right)^{L_{\mathrm{MU}}}\right]}{2}\left(\frac{\breve{\alpha}_1}{4}\xi_i\gamma+1\right)^{-N_iL_{\mathrm{MU}}}$$
$$=\left[\sqrt[N_iL_{\mathrm{MU}}]{\frac{2}{\prod_{l=1}^{N_i}\left(\frac{\breve{\alpha}_1}{\breve{\alpha}_l}\right)^{L_{\mathrm{MU}}}}}\frac{\breve{\alpha}_1}{4}\xi_i\gamma+\sqrt[N_iL_{\mathrm{MU}}]{\frac{2}{\prod_{l=1}^{N_i}\left(\frac{\breve{\alpha}_1}{\breve{\alpha}_l}\right)^{L_{\mathrm{MU}}}}}\right]^{-N_iL_{\mathrm{MU}}}$$
$$\approx\gamma^{-N_iL_{\mathrm{MU}}}\left[\frac{\prod_{l=1}^{N_i}\left(\frac{\breve{\alpha}_1}{\breve{\alpha}_l}\right)^{L_{\mathrm{MU}}}}{2}\left(\frac{\breve{\alpha}_1}{4}\xi_i\right)^{-N_iL_{\mathrm{MU}}}\right]+\mathrm{o}\left(\gamma^{-N_iL_{\mathrm{MU}}}\right)$$
$$\leq\left[\frac{\breve{\alpha}_1\sqrt[N_iL_{\mathrm{MU}}]{2}}{4}\xi_i\gamma\right]^{-N_iL_{\mathrm{MU}}}+\mathrm{o}\left(\gamma^{-N_iL_{\mathrm{MU}}}\right). \tag{32}$$

the i-th user is given in (29) at the top of this page. Thus, the evaluation of the ABEP of the i-th user follows from (8), by using (29).

In addition to the performance of each user, the whole system performance is of interest. A metric that is able to evaluate the whole system performance is the system ABEP. Assuming that the detection process at each user is performed independently, the system ABEP is expressed as:

$$P_{\mathrm{bit}}^{\mathrm{System}}(\gamma)=\frac{1}{N_u}\sum_{i}^{Nu}P_{\mathrm{bit}}^{i}(\gamma_i). \tag{30}$$

An upper bound of (30) can be obtained by using the upper bound of the PEP, $P_{\mathrm{bit}}^{i}(\gamma_i)$, of each user given in (8).

V. Analysis of Diversity Order and Coding Gain

In the high SNR regime, the user and system performance can be characterized in terms of diversity order and coding gain. In [42], the diversity order and coding gain are obtained from the metric of Symbol Error Rate (SER). Therefore, in order to be perfectly aligned with [42], the focus in this section is on the metric of SER. We start by analyzing these performance measures for the i-th user, $i=1,\ldots,N_u$ and then generalize the analysis from the system standpoint based on (30).

In order to compute the diversity order and coding gain of the i-th user, a high SNR approximation for the PEPs of the i-th user is needed. By using mathematical steps similar to Section IV and based on the Chernoff bound of the Q-function, $Q(x)\leq\frac{1}{2}e^{-\frac{x^2}{2}}$, the PEP of the i-th user can be bounded as:

$$P_{\mathrm{e}}^{i}(\mathbf{x}_i\to\hat{\mathbf{x}}_i,\gamma)\leq\frac{\left[\prod_{l=1}^{N_i}\left(\frac{\breve{\alpha}_1}{\alpha_l}\right)^{L_{\mathrm{MU}}}\right]}{2}\left(\frac{\breve{\alpha}_1}{4}\xi_i\gamma+1\right)^{-N_iL_{\mathrm{MU}}}$$
$$\times\sum_{k=0}^{+\infty}\breve{\delta}_k\left(\frac{\breve{\alpha}_1}{4}\xi_i\gamma+1\right)^{-k}. \tag{31}$$

From (31), a high SNR ($\gamma\to+\infty$) approximation of the PEP of the i-th user can be obtained as given below. If the SNR approaches to infinity, only the smallest value of the exponent k in (31) needs to be considered, which is equal to one. In this way, using the previous simplification, (31) can be further approximated as shown in (32) at the top of this page. Note that the last step in (32) follows from the inequality:

$$\prod_{l=1}^{N_i}\left(\frac{\breve{\alpha}_1}{\breve{\alpha}_l}\right)^{L_{\mathrm{MU}}}\leq 1, \tag{33}$$

which holds because $\breve{\alpha}_l$, $l=1,\ldots,N_i$, are the eigenvalues of $\breve{\mathbf{A}}$ in (22) in ascending order.

In this case, a high SNR approximation of the SER of the i-th user is obtained as follows:

$$\mathrm{SER}_i^{+\infty}\lessapprox\frac{1}{|\mathcal{B}|}\sum_{\mathbf{x}_i}\sum_{\substack{\hat{\mathbf{x}}_i\\ \hat{\mathbf{x}}_i\neq\mathbf{x}_i}}P_i^{+\infty}(\mathbf{x}_i\to\hat{\mathbf{x}}_i,\gamma). \tag{34}$$

In (34), $\mathcal{B}$ denotes the set of all possible transmitted symbol vectors to a generic user.

It can be observed that the high SNR approximation of the SER in (34) is a linear combination of $P_i^{+\infty}(\mathbf{x}_i \to \hat{\mathbf{x}}_i, \gamma)$, as given in (32), for all possible pairs of $\mathbf{x}_i$ and $\hat{\mathbf{x}}_i$. Therefore, as $\gamma \to +\infty$, the slope of (34) is determined by the smallest exponent of γ in (32), i.e. $N_i L_{\text{MU}}$. The smallest value of $N_i L_{\text{MU}}$ occurs when $N_i = 1$. In fact, the dominant addends of (34) are those for which $N_i = 1$. Therefore, the high SNR approximation of the SER in (34) can be further approximated by considering only these dominant addends. In addition, the careful inspection of (32) shows that the matrix $\breve{\mathbf{A}}$ in (22) reduces to a scalar if $N_i = 1$. This implies $\breve{\alpha}_1 = b_1$, where b_1 is given in (15).

With this simplification at hand, a more insightful approximation of (34) can be obtained. More specifically, from (32), (34) can be expressed as:

$$\text{SER}_i^{+\infty} \lessapprox \frac{1}{|\mathcal{B}|} \sum_{\mathbf{x}_i} \sum_{n=1}^{N_s} \sum_{x \in \mathcal{M}} \sum_{\substack{\hat{x} \in \mathcal{M} \\ \hat{x} \neq x}} \left[\frac{|x - \hat{x}|^2 \sqrt[L_{\text{MU}}]{2}}{4} \xi_i \gamma \right]^{-L_{\text{MU}}} + \sum_{k=2}^{N_s} \text{o}\left(\gamma^{-k L_{\text{MU}}}\right). \tag{35}$$

By using a line of thought similar to [35, Chapter 5.2.9], an upper bound for (35) is obtained by retaining only the minimum distance, denoted by d_{min}, between every pair of the constellation points $\{x, \hat{x}\} \in \mathcal{M}$. By doing so, after some algebraic manipulations of the summations in (35), the following high SNR approximation of the SER of the i-th user is obtained:

$$\text{SER}_i^{+\infty} \lessapprox \left[\frac{d_{\text{min}}^2}{4} \sqrt[L_{\text{MU}}]{\frac{2}{N_s \left(M^2 - M\right)}} \xi_i \gamma \right]^{-L_{\text{MU}}} + \sum_{k=2}^{N_s} \text{o}\left(\gamma^{-k L_{\text{MU}}}\right). \tag{36}$$

The bound in (36) may be loose for high values of the constellation order M [35, Chapter 5.2.9]. However, it is conveniently formulated for providing insightful information on the achievable diversity order and coding gain. If M is large, if needed, a tighter bound may be obtained by following the guidelines in [35, Chapter 5.2.9].

Based on the definitions of the diversity order and coding gain available in [42], the inspection of (36) reveals that the diversity order of the i-th user is:

$$d_i = L_{\text{MU}}, \tag{37}$$

and that the corresponding coding gain is:

$$c_i = \frac{d_{\text{min}}^2}{4} \sqrt[L_{\text{MU}}]{\frac{2}{N_s \left(M^2 - M\right)}} \xi_i. \tag{38}$$

Since $L_{\text{MU}} = N_t - N_u N_r + 1$, from (37) it follows that the diversity order of the i-user does not depend on large-scale channel effect, but only on the system size (the number of transmit antennas N_r, the number of users N_u, and the number of receive antennas per user N_r). In contrast, (38) shows that the coding gain of the i-th user depends on the system size, the number of parallel data streams N_s, the constellation size $\mathcal{M}$ (via d_{min} and M), and ξ_i which represents the large-scale channel effect. Here, it is indirectly assumed that the large-scale channel effect is deterministic.

From the system-level standpoint, the diversity order can be computed by approximating the system SER for high SNR ($\gamma \to +\infty$) as follows:

$$\begin{aligned} \text{SER}_{\text{System}}^{+\infty} &= \frac{1}{N_u} \sum_{i=1}^{N_u} \text{SER}_i^{+\infty} \\ &\lessapprox \frac{1}{N_u} \sum_{i=1}^{N_u} \left[\frac{d_{\text{min}}^2}{4} \sqrt[L_{\text{MU}}]{\frac{2}{N_s \left(M^2 - M\right)}} \xi_i \gamma \right]^{-L_{\text{MU}}} \\ &\quad + \sum_{k=2}^{N_s} \text{o}\left(\gamma^{-k L_{\text{MU}}}\right), \end{aligned} \tag{39}$$

where, the last step in (39) exploits the high SNR approximation in (36). In order to express (39) in a convenient form that explicitly provides information on the diversity order and coding gain, an upper bound based on the smallest value of ξ_i, $i = 1, \ldots, N_u$ is used. More specifically, the following holds:

$$\text{SER}_{\text{System}}^{+\infty} \lessapprox \left[\frac{d_{\text{min}}^2}{4} \sqrt[L_{\text{MU}}]{\frac{2}{N_s \left(M^2 - M\right)}} \xi_{\text{min}} \gamma \right]^{-L_{\text{MU}}} + \sum_{k=2}^{N_s} \text{o}\left(\gamma^{-k L_{\text{MU}}}\right), \tag{40}$$

where, $\xi_{\text{min}} = \min\left(\xi_1, \ldots, \xi_{N_u}\right)$. From (40) and [42], it follows that the system diversity order is:

$$d_{\text{System}} = L_{\text{MU}}, \tag{41}$$

and the corresponding coding gain is:

$$c_{\text{System}} = \frac{d_{\text{min}}^2}{4} \sqrt[L_{\text{MU}}]{\frac{2}{N_s \left(M^2 - M\right)}} \xi_{\text{min}}. \tag{42}$$

Comparing (41) and (42) with (37) and (38), respectively, we conclude that the diversity order and the coding gain from the user and system standpoints are the same. The main difference is that the system-level coding gain in (42) is dominated by the large-scale channel effect of the user having the weakest channel, i.e., the smallest value of ξ_i, $i = 1, \ldots, N_u$.

Based on the obtained expressions of the diversity order and coding gain, the proposed transmission scheme can be compared against the conventional MIMO broadcast channel. To this end, it is worth noting that the proposed mathematical framework is directly applicable to the conventional MIMO broadcast channel by simply setting $N_s = N_r$. Therefore, (37), (38), (41), and (42) can be directly used for comparing MU MSR-SM and the conventional MIMO broadcast channel.

From (37) and (41), in particular, we conclude that the diversity order is independent of N_s. As a result, both schemes achieve the same diversity order. As for the conventional MIMO broadcast channel, this conclusion is in agreement with the results available in [43]. This further validates the

correctness of our mathematical framework.

The comparison of (38) and (42), on the other hand, brings to our attention that the coding gain depends on N_s. Therefore, the coding gain of MU MSR-SM, where, in general, $N_s < N_r$ holds, is different from the coding gain of the conventional MIMO broadcast channel, where $N_s = N_r$. Since both schemes offer the same diversity order, the scheme providing the highest coding gain also results in the lowest BER. Hence, the superiority of a scheme compared to the other can be assessed by a direct inspection of the following coding gain ratio:

$$\lambda_i = \frac{c_i^{\mathrm{MSR-SM}}}{c_i^{\mathrm{SMX}}} = \left(\frac{d_{\min}^{\mathrm{MSR-SM}}}{d_{\min}^{\mathrm{SMX}}}\right)^2 \sqrt[L_{\mathrm{MU}}]{\frac{N_r\left(M_{\mathrm{SMX}}^2 - M_{\mathrm{SMX}}\right)}{N_s\left(M_{\mathrm{MSR-SM}}^2 - M_{\mathrm{MSR-SM}}\right)}}. \quad (43)$$

By appropriately choosing the constellation orders $M_{\mathrm{MSR-SM}}$ and M_{SMX} for MU MSR-SM and for the conventional MIMO broadcast channel, respectively, the same spectral efficiency can be guaranteed. In (43), the coding gain of the i-th user of MU MSR-SM is denoted by $c_i^{\mathrm{MSR-SM}}$ and the coding gain of the same user in the conventional MIMO broadcast channel is denoted by c_i^{SMX}. Furthermore, $d_{\min}^{\mathrm{MSR-SM}}$ and $d_{\min}^{\mathrm{SMX}}$ denote the minimum distance between every pair of points of the adopted signal constellations for MU MSR-SM and for the conventional MIMO broadcast channel, respectively. From (43), it follows that MU MSR-SM performs better than the conventional MIMO broadcast channel if $\lambda_i > 1$.

If $M_{\mathrm{MSR-SM}} = M_{\mathrm{SMX}}$, a direct inspection of (43) reveals that $\lambda_i > 1$ and that it increases as N_s decreases. This is also supported by the fact that as N_s is reduced, $d_{\min}^{\mathrm{MSR-SM}}$ is increased. This happens because, for a fair comparison, the power of $\mathbf{x}$ and $\mathbf{x}_i$ should be irrespective of N_s . Therefore, the distances between the points of a deployed constellation are increased as N_s is reduced. As a result, in this case, MU MSR-SM outperforms the conventional MIMO broadcast channel.

In general, however, it holds that $M_{\mathrm{MSR-SM}} \neq M_{\mathrm{SMX}}$. In this case, usually, MU MSR-SM deploys a constellation of higher order in order to achieve the same spectral efficiency as the conventional MIMO broadcast channel. Therefore, the minimum distance, $d_{\min}^{\mathrm{MSR-SM}}$, between every pair of points of the deployed constellation is decreased. Hence, although the decrease of N_s have a positive effect on the increase of the value of λ_i, the combined effect of increasing the constellation order and consequently reducing $d_{\min}^{\mathrm{MSR-SM}}$ may result in lower values of λ_i.

In this case, a direct analysis of (43) is more difficult. The ratio λ_i can, however, be numerically computed. Table I provides typical values of λ_i in dB scale, by assuming the same spectral efficiency for both schemes. The inspection of Table I shows that, for a group of system setups ($N_t = 20$, $N_r = 4$, $N_u = 4$, $N_s = 3$; $N_t = 20$, $N_r = 5$, $N_u = 4$, $N_s = 4$; and $N_t = 16$, $N_r = 4$, $N_u = 4$, $N_s = 3$), MU MSR-SM provides a higher coding gain than the conventional MIMO broadcast channel. More specifically, the coding gain is in the range between 1 and 2.49 dB. On the other hand, for the rest of the system setups, the conventional MIMO broadcast channel provides a higher coding gain. As a result, (43) can be used for the system optimization and for ensuring that MSR-SM is superior to the state-of-the-art. Note that, as shown in Section II, smaller values of N_s result in lower computational complexity at the transmitter. Therefore, in terms of coding gain and complexity at the transmitter, the optimal way for selecting the value of N_s is to find the smallest one for which it holds that $\lambda_i \geq 1$.

TABLE I
CODING GAIN OF MU MSR-SM, BASED ON (43), WITH RESPECT TO THE CONVENTIONAL MIMO BROADCAST CHANNEL ($N_s = N_r$).

System Configuration	N_s	k_{user} (bpsp)	λ_i (in dB)
$N_t = 20$, $N_r = 4$, $N_u = 4$	3	8	1.49
$N_t = 20$, $N_r = 4$, $N_u = 4$	2	8	-2.49
$N_t = 20$, $N_r = 4$, $N_u = 4$	1	8	-11.05
$N_t = 20$, $N_r = 5$, $N_u = 4$	4	10	1.93
$N_t = 16$, $N_r = 4$, $N_u = 4$	3	8	2.49
$N_t = 10$, $N_r = 2$, $N_u = 4$	1	8	-2.98

VI. Simulation Results and Discussion

The objective of this section is twofold. First, to validate the theoretical results of Sections IV and V using simulation results. Second, to provide a performance comparison between MU MSR-SM and the benchmark MIMO broadcast channel. In this latter case, in particular, $N_s = N_r$ is assumed and no SM is used. More specifically, the benchmark system conveys information to the N_u remote users by establishing N_r parallel and non-interfering data streams to each one of them. In all studied scenarios, the number of users is equal to $N_u = 4$. As described in Section II, the wireless channel of the i-th user, $i = 1, \ldots, N_u$, is generated following a complex Gaussian distribution $\left(\check{\mathbf{H}}_i \sim \mathcal{CN}(\mathbf{0}, \xi_i \mathbf{I})\right)$. In more detail, ξ_i, $i = 1, \ldots, N_u$, is set equal to 1, 0.75, 0.5, and 0.25, for user 1, 2, 3, and 4, respectively. This choice allows us to demonstrate how the large-scale channel effect (system topology) affects the performance of different users. Values of ξ_i which are close to one model strong channels, while as ξ_i is reduced and approaches zero, less strong channels are modeled. Note that, if the effect of shadowing is not considered as it is a common assumption in the literature [34, 38, 39, 44–47], the value of ξ_i is solely determined by the transmission distance (pathloss). More specifically, for the i-th user and at a normalized distance $r_i = \bar{r}_i / r_0$, where $\bar{r}_i$ is the transmission distance and r_0 is a given reference distance, the value of ξ_i is given as, $\xi_i = 1/r_i^{\alpha}$. Here, $\alpha \geq 2$ is the pathloss exponent. Therefore, for $\alpha = 2$, the previous values of ξ_i correspond to the following normalized distances of 1, 1.154, 1.142, and 2, respectively. For a fair comparison, the M-ary constellations of both schemes are normalized such that $\mathrm{E}_{\mathbf{x}}[\mathbf{x}] = 1$ and $\mathrm{E}_{\mathbf{x}_i}[\mathbf{x}_i] = \frac{1}{N_u}$. So, the transmit SNR of the whole system is $\gamma = \frac{1}{\sigma_{\mathbf{w}}^2}$.

A. Validation of (21)

Section IV provides the ABEP of MU MSR-SM by using the PDF of z_i given in (21). The derivation of (21) is based

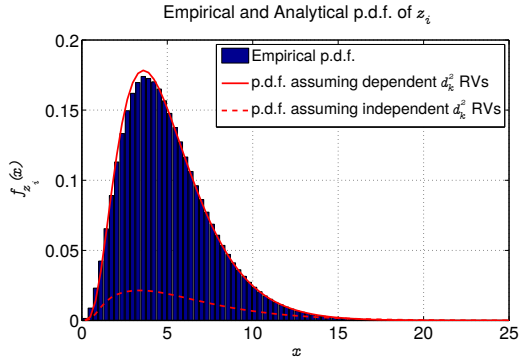


Fig. 1. Illustration of empirical and analytical p.d.f of (17) by assuming that: i) the RVs d_k, $k = 1, \ldots, 2$ are statistically dependent and ii) they are independent. Setup: $\mathbf{H} \sim \mathcal{CN}(\mathbf{0}_{2\times 4}, \mathbf{I}_{2\times 4})$; and ii) $b_1 = 0.5$ and $b_1 = 1.2$.

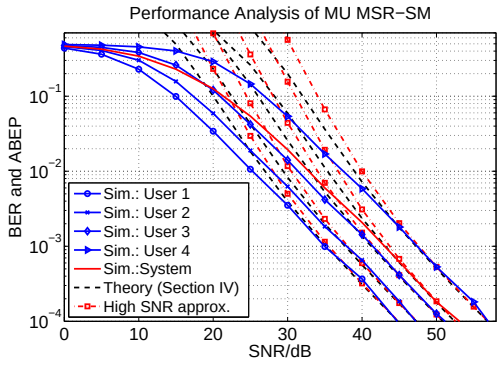


Fig. 2. Performance analysis of MU R-SM ($N_s = 1$) for four users, when ZF with P-CSIT is employed: simulation results vs. the bounds in Section IV. The high SNR approximation of the ABEP is calculated using the PEP given in (32). Setup: $N_t = 16$, $N_r = 4$, ξ_i, $i = 1, \ldots, N_u$, takes the values 1, 0.75, 0.5, and 0.25 for the user 1, 2, 3, and 4, respectively.

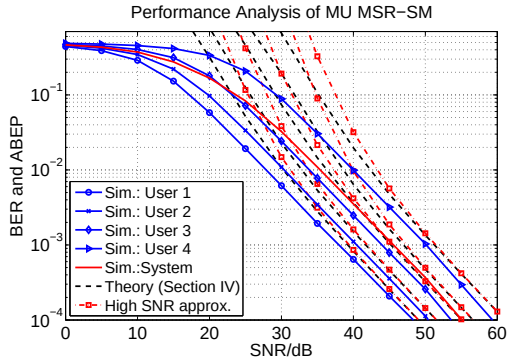


Fig. 3. Performance analysis of MU MSR-SM ($N_s = 2$) for four users, when ZF with P-CSIT is employed: simulation results vs. the bounds in Section IV. The high SNR approximation of the ABEP is calculated using the PEP given in (32). Setup: $N_t = 16$, $N_r = 4$, ξ_i, $i = 1, \ldots, N_u$, takes the values 1, 0.75, 0.5, and 0.25 for the user 1, 2, 3, and 4, respectively.

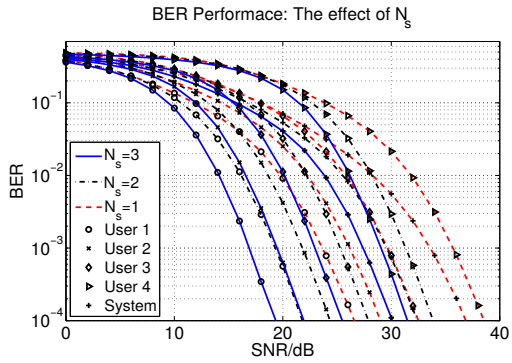
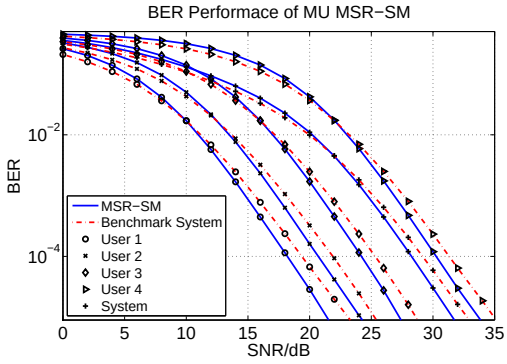


Fig. 4. BER performance of MU MSR-SM as a function of N_s. Setup: $N_t = 20$, $N_r = 4$, $N_u = 4$. The spectral efficiency is 8 bpcu.

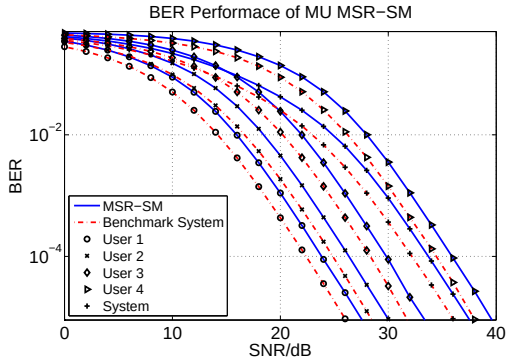
on the fact that the RVs d_k^2, $k = 1, \ldots, N_r$, are statistically correlated. In order to confirm this, Fig. 1 illustrates the empirical PDF of (17) and compares it against its analytical expression in (21). In addition, Fig. 1 shows the analytical PDF of (17) under the assumption that d_k^2, $k = 1, \ldots, N_r$, are assumed to be statistically independent RVs, as usually considered in the literature for mathematical tractability. If this assumption was valid, the PDF of (17) could be directly obtained by using the result from [40, Theorem 1]. From Fig. 1, we observe that the theoretical PDF of (17) perfectly matches its empirical PDF. In contrast, when the RVs d_k^2, $k = 1, \ldots, N_r$ are assumed to be independent, the obtained PDF from [40, Theorem 1] deviates from the empirical results.

B. Validation of the Theoretical Analysis

The upper bounds derived in Section IV are compared against Monte Carlo simulations in Figs. 2 and 3. Furthermore, the same figures illustrate the upper bounds of the ABEP for the system and for each user when the high SNR approximation of the PEP in (32) is used. Note that Figs. 2 and 3 present the BER of the proposed architecture in very high SNRs solely for validating the theoretical framework of this paper. The inspection of Figs. 2 and 3 indicates that the analytical bounds of the ABEP are tight in the high SNR, both for (29) and (32). More specifically, in the high SNR, the analytical results can be considered as an excellent approximation of the simulation results. In contrast, in the low SNR, there is a small difference between the theoretical and simulation results. However, this is a well known phenomenon that originates from using union bound methods [35]. Finally, the diversity order and coding gain analysis of Section V is also verified from Figs. 2 and 3. In more detail, the slope of the simulated BER curves of each user is $L_{\text{MU}} = N_t - N_u N_r + 1$. In addition, the simulated curves show that the behavior of the coding gain of the i-th user depends on ξ_i. Similarly, the simulated curves show that the system coding gain is dominated by $\xi_{\min}$. These conclusions are in perfect agreement with the analysis presented in Section V.

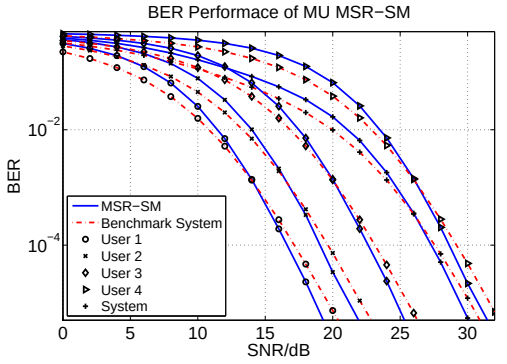


(a) The spectral efficiency is 2 bpcu, with $N_s = 1$.

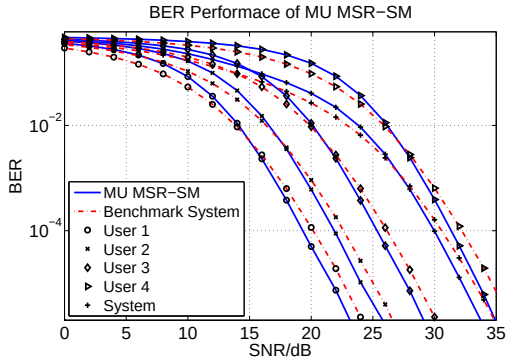


(b) The spectral efficiency is 4 bpcu, with $N_s = 1$.

Fig. 5. BER performance of MU MSR-SM versus benchmark system (conventional MIMO broadcast channel with SMX). Setup: $N_t = 10$, $N_r = 2$, $N_u = 4$.



(a) The spectral efficiency is 4 bpcu, with $N_s = 2$.



(b) The spectral efficiency is 8 bpcu, with $N_s = 3$.

Fig. 6. BER performance of MU MSR-SM versus benchmark system (conventional MIMO broadcast channel with SMX). Setup: $N_t = 20$, $N_r = 4$, $N_u = 4$.

C. Analysis of the Impact of N_s

Figure 4 shows the BER of MU MSR-SM as a function of N_s. It shows, in particular, that the BER of each user gets better as N_s increases. This happens because higher values of N_s require a lower modulation order of Quadrature Amplitude Modulation (QAM) in order to achieve the same spectral efficiency. Table I, however, shows that the optimal value of N_s is not necessarily equal to N_r, i.e., the conventional MIMO broadcast channel.

D. BER Comparison with the Conventional Broadcast Channel

Figures 5 and 6 show the BER of MU MSR-SM and the conventional MIMO broadcast channel (benchmark system) for different system setups. As shown in Figs. 5(a), 6(a), and 6(b), in the low SNR, conventional MIMO offers a slightly better BER than the new scheme. In the high SNR, on the other hand, the new architecture outperforms the benchmark system. More specifically, in Figs. 5(a), 6(a), and 6(a), and at BER=10^{-4}, MU MSR-SM provides a gain of 1.4, 0.8, and 1 dB, respectively. In contrast, in Fig. 6(b), MU MSR-SM is outperformed by the benchmark system. However, even in this case, MU MSR-SM retains its complexity benefits at the transmitter. These findings are in agreement with the mathematical analysis of the ratio of the coding gains in (43).

E. Energy Efficiency Comparison with the Conventional Broadcast Channel

In this section, the energy efficiency of MU MSR-SM is studied using the Relative Average Energy Reduction (RAER) performance metric, which is defined as follows:

$$\text{RAER}[\%] = \left[1 - 10^{-\frac{\Delta_{\text{SNR}}}{10}}\right] \times 100\%, \qquad (44)$$

where, Δ_{SNR} denotes the SNR difference (in dB) between MU MSR-SM and the conventional MIMO broadcast channel for a given BER. It is worth mentioning that only the energy consumption for the RF power transmission is considered. Both architectures, in fact, have almost the same circuits

TABLE II
RAER OF MU MSR-SM WITH RESPECT TO THE CONVENTIONAL MIMO BROADCAST CHANNEL.

Configuration	k_{system} (bpsp)	Target BER	RAER [%]
$N_t = 20$, $N_r = 4$, $N_u = 4$	32	10^{-1}	-20%
$N_t = 20$, $N_r = 4$, $N_u = 4$	32	10^{-2}	1%
$N_t = 20$, $N_r = 4$, $N_u = 4$	32	10^{-4}	18%

energy consumption, since they employ the same number of RF front-ends.

Table II presents the system RAER of MU MSR-SM with respect to the conventional MIMO broadcast channel, by assuming the same system setup as in Fig. 6(a). Both schemes provide the same BER and the same spectral efficiency. If BER=10^{-1}, the conventional MIMO broadcast channel is more energy efficient. For practical values of the BER less than 10^{-2}, on the other hand, MU MSR-SM becomes more energy efficient. For example, an energy efficiency gain of 18% is achieved by MU MSR-SM at BER=10^{-4}.

VII. CONCLUSIONS

In this paper, the incorporation of MSR-SM for application to the MIMO broadcast channel is introduced and its BER performance is mathematically studied. More specifically, based on the union bound technique, an accurate mathematical framework for its performance evaluation is proposed and discussed. From this framework, it is proved that MSR-SM provides the same diversity order as the conventional MIMO broadcast channel, while offering a better coding gain in the high SNR regime. Also, this performance gain is achieved with a reduction of the complexity of the transmitter. Numerical simulations are shown in order to substantiate the gain predicted by the analysis. As a result, MSR-SM is shown to be a promising transmission scheme for the MIMO broadcast channel.

REFERENCES

[1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 6, pp. 311–335, 1998.

[2] E. Telatar, "Capacity of multi-antenna gaussian channels," *European Trans. on Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov. / Dec. 1999.

[3] M. Di Renzo, H. Haas, A. Ghrayeb, S. Sugiura, and L. Hanzo, "Spatial modulation for generalized MIMO: Challenges, opportunities, and implementation," *Proc. IEEE*, vol. 102, no. 1, pp. 56–103, Jan 2014.

[4] M. Di Renzo, H. Haas, and P. M. Grant, "Spatial modulation for multiple-antenna wireless systems: A survey," *IEEE Commun. Mag.*, vol. 49, no. 12, pp. 182–191, December 2011.

[5] P. Yang, M. Di Renzo, Y. Xiao, S. Li, and L. Hanzo, "Design guidelines for spatial modulation," *IEEE Commun. Surveys Tut.*, vol. 17, no. 1, pp. 6–26, Firstquarter 2015.

[6] R. Mesleh, H. Haas, S. Sinanović, C. W. Ahn, and S. Yun, "Spatial modulation," *IEEE Trans. on Veh. Tech.*, vol. 57, no. 4, pp. 2228 – 2241, Jul. 2008.

[7] J. Jeganathan, A. Ghrayeb, and L. Szczecinski, "Spatial modulation: Optimal detection and performance analysis," *IEEE Commun. Lett.*, vol. 12, no. 8, pp. 545–547, 2008.

[8] A. Stavridis, S. Sinanović, M. D. Renzo., and H. Haas, "Energy evaluation of spatial modulation at a multi-antenna base station," in *Proc. of the 78th IEEE Veh. Tech. Conf. (VTC)*, Las Vegas, USA, Sep. 2–5, 2013.

[9] J. Jeganathan, A. Ghrayeb, L. Szczecinski, and A. Ceron, "Space shift keying modulation for MIMO channelsmeh," *IEEE Trans. on Wireless Commun.*, vol. 8, no. 7, pp. 3692–3703, Jul. 2009.

[10] J. Jeganathan, A. Ghrayeb, and L. Szczecinski, "Generalized Space Shift Keying Modulation for MIMO Channels," in *Proc. IEEE 19th Intern. Symp. on Personal, Indoor and Mobile Radio Commun. PIMRC 2008*, Cannes, France, 15–18 Sep. 2008, pp. 1–5.

[11] M. Raed, I. Salama, A. Hadi, and A. Mansour, "Performance analysis of space shift keying (SSK) modulation with multiple cooperative relays," *EURASIP J. on Adv. in Signal Process.*, 2012.

[12] M. Wen, X. Cheng, H. Poor, and B. Jiao, "Use of SSK modulation in two-way amplify-and-forward relaying," *IEEE Trans. on Veh. Techn.*, vol. 63, no. 3, pp. 1498–1504, March 2014.

[13] P. Som and A. Chockalingam, "BER analysis of space shift keying in cooperative multi-hop multi-branch DF relaying," in *Proc. of the 78th IEEE Veh. Tech. Conf. (VTC)*, Las Vegas, USA, Sep. 2–5, 2013.

[14] ——, "Performance analysis of space-shift keying in decode-and-forward multihop MIMO networks," *IEEE Trans. on Veh. Techn.*, vol. 64, no. 1, pp. 132–146, Jan 2015.

[15] M. Maleki, H. Bahrami, S. Beygi, M. Kafashan, and N. Tran, "Space modulation with csi: Constellation design and performance evaluation," *IEEE Trans. on Vehicular Techn.*, vol. 62, no. 4, pp. 1623–1634, May 2013.

[16] P. Yang, Y. Xiao, Y. Yu, and S. Li, "Adaptive Spatial Modulation for Wireless MIMO Transmission Systems," *IEEE Commun. Lett.*, vol. 15, no. 6, pp. 602–604, June 2011.

[17] P. Yang, Y. Xiao, Y. Yu, L. Li, Q. Tang, and S. Li, "Simplified adaptive spatial modulation for limited-feedback MIMO systems," *IEEE Trans. on Veh. Technol.*, vol. 62, no. 6, pp. 2656–2666, July 2013.

[18] P. Yang, Y. Xiao, B. Zhang, S. Li, M. El-Hajjar, and L. Hanzo, "Power allocation-aided spatial modulation for limited-feedback MIMO systems," *IEEE Trans. on Veh. Technol.*, vol. 64, no. 5, pp. 2198–2204, May 2015.

[19] M. Veedu, C. Murthy, and L. Hanzo, "Single-RF spatial modulation relying on finite-rate phase-only feedback: Design and analysis," *IEEE Trans. on Veh. Technol.*, vol. PP, no. 99, pp. 1–1, 2015.

[20] S. Sugiura, S. Chen, and L. Hanzo, "Generalized space-time shift keying designed for flexible diversity-, multiplexing- and complexity-tradeoffs," *IEEE Trans. on Wireless Commun.*, vol. 10, no. 4, pp. 1144 –1153, Apr. 2011.

[21] N. Serafimovski, A. Younis, R. Mesleh, P. Chambers, M. D. Renzo, C.-X. Wang, P. M. Grant, M. A. Beach, and H. Haas, "Practical Implementation of Spatial Modulation," *IEEE Trans. on Veh. Tech.*, vol. 62, no. 9, pp. 4511–4523, 2013.

[22] A. Younis, W. Thompson, M. D. Renzo, C.-X. Wang, M. A. Beach, H. Haas, and P. M. Grant, "Performance of spatial modulation using measured real-world channels," in *Proc. of the 78th IEEE Veh. Tech. Conf. (VTC)*, Las Vegas, USA, Sep. 2–5 2013.

[23] J. Zhang, Y. Wang, L. Ding, and N. Zhang, "Bit error probability of spatial modulation over measured indoor channels," *IEEE Trans. on Wireless Commun.*, vol. 13, no. 3, pp. 1380–1387, March 2014.

[24] L.-L. Yang, "Transmitter preprocessing aided spatial modulation for multiple-input multiple-output systems," in *Proc. of 73rd IEEE Veh.Techn. Conf. (VTC Spring)*, May 2011, pp. 1 –5.

[25] R. Zhang, L.-L. Yang, and L. Hanzo, "Generalised pre-coding aided spatial modulation," *IEEE Trans. on Wireless Commun.*, vol. 12, no. 11, pp. 5434–5443, November 2013.

[26] ——, "Error probability and capacity analysis of generalised pre-coding aided spatial modulation," *IEEE Trans. on Wireless Commun.*, vol. PP, no. 99, pp. 1–1, 2014.

[27] A. Stavridis, S. Sinanović, M. Di Renzo, and H. Haas, "Transmit precoding for receive spatial modulation using imperfect channel knowledge," in *Proc. of 75th IEEE Veh. Techn. Conf. (VTC Spring)*, 2012.

[28] A. Stavridis, D. Basnayaka, M. Di Renzo, and H. Haas, "Average bit error probability of receive-spatial modulation using zero-forcing precoding," in *IEEE 19th Int. Workshop on Computer Aided Modeling and Des. of Commun. Links and Netw. (CAMAD)*, 2014.

[29] A. Stavridis, D. Basnayaka, S. Sinanovic, M. Di Renzo, and H. Haas, "A virtual MIMO dual-hop architecture based on hybrid spatial modulation," *IEEE Trans. on Commun.*, vol. 62, no. 9, pp. 3161–3179, Sept 2014.

[30] M. Kountouris and J. Andrews, "Downlink SDMA with limited feedback in interference-limited wireless networks," *IEEE Trans. on Wireless Commun.*, vol. 11, no. 8, pp. 2730–2741, August 2012.

[31] M. Matthaiou, N. Chatzidiamantis, G. Karagiannidis, and J. Nossek, "ZF detectors over correlated K fading MIMO channels," *IEEE Trans. on Commun,.*, vol. 59, no. 6, pp. 1591–1603, June 2011.

[32] S. Narayanan, M. J. Chaudhary, A. Stavridis, R. Di Renzo, F. Graziosi, and H. Haas, "Multi-user spatial modulation MIMO," in *Proc. of IEEE*

Wireless Communications and Networking Conference (WCNC), Apr. 6–9, 2014.

[33] X. Li, Y. Zhang, L. Xiao, X. Xu, and J. Wang, "A novel precoding scheme for downlink multi-user spatial modulation system," in *Proc. IEEE 24th Int. Symp. Pers. Indoor and Mobile Radio Commun. (PIMRC)*, Sept 2013, pp. 1361–1365.

[34] A. Wiesel, Y. Eldar, and S. Shamai, "Zero-Forcing Precoding and Generalized Inverses," *IEEE Trans. on Signal Process.*, vol. 56, no. 9, pp. 4409 –4418, Sept. 2008.

[35] J. G. Proakis, *Digital Communications*, 4th ed. New York, NY, USA: McGraw–Hill, 2000.

[36] M. Chiani, D. Dardari, and M. K. Simon, "New exponential bounds and approximations for the computation of error probability in fading channels," *IEEE Trans. on Wireless Commun.*, vol. 2, no. 4, pp. 840–845, July 2003.

[37] X. Shao, J. Yuan, and Y. Shao, "Error performance analysis of linear zero forcing and MMSE precoders for MIMO broadcast channels," *IET Commun.*, vol. 1, no. 5, pp. 1067–1074, Oct 2007.

[38] C.-J. Chen and L.-C. Wang, "Performance analysis of scheduling in multiuser MIMO systems with zero-forcing receivers," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1435–1445, September 2007.

[39] D. Gore, R. Heath, and A. Paulraj, "Transmit selection in spatial multiplexing systems," *IEEE Commun. Lett.*, vol. 6, no. 11, pp. 491–493, Nov 2002.

[40] M.-S. Alouini, A. Abdi, and M. Kaveh, "Sum of gamma variates and performance of wireless communication systems over Nakagami-fading channels," *IEEE Trans. on Veh. Technol.*, vol. 50, no. 6, pp. 1471–1480, Nov 2001.

[41] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed., A. Jeffrey and D. Zwillinger, Eds. Academic Press, Mar. 2007, ISBN-10: 0123736374.

[42] Z. Wang and G. Giannakis, "A simple and general parameterization quantifying performance in fading channels," *IEEE Trans. on Commun.*, vol. 51, no. 8, pp. 1389–1398, Aug 2003.

[43] A. Mehana and A. Nosratinia, "Diversity of MIMO linear precoding," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 1019–1038, Feb 2014.

[44] W. Guan and K. Liu, "Diversity analysis of analog network coding with multi-user interferences," *IEEE Trans. Wireless Commun.*, vol. 12, no. 2, pp. 668–679, February 2013.

[45] H. Dhillon, R. Ganti, F. Baccelli, and J. Andrews, "Modeling and analysis of k-tier downlink heterogeneous cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 3, pp. 550 –560, april 2012.

[46] M. Di Renzo and W. Lu, "Stochastic geometry modeling and performance evaluation of MIMO cellular networks using the equivalent-in-distribution (EiD)-based approach," *EEE Trans. on Commun.*, vol. 63, no. 3, pp. 977–996, March 2015.

[47] K. Ntontin, M. Di Renzo, A. Perez-Neira, and C. Verikoukis, "Analog network coding in the multiple access relay channel: Error rate analysis and optimal power allocation," *IEEE Trans. on Wireless Commun.*, vol. 14, no. 6, pp. 3015–3032, June 2015.

Athanasios Stavridis holds the Diploma degree (M.Eng. equivalent) in computer engineering and informatics and the M.Sc. degree in signal processing and communication systems from the University of Patras, Patras, Greece, since 2006 and 2008, respectively, and the PhD degree in wireless communication from The University of Edinburgh, Edinburgh, U.K, since 2015. Since 2014, he has been a Research Associate with The University of Edinburgh, working on RF and visible light communication.

In 2008, Athanasios was awarded with a scholarship from the University of Patras upon the completion of his M.Sc. studies. During the period of June 2011–June 2014, he was a Marie Curie Early-Stage-Researcher at The University of Edinburgh, Edinburgh, U.K. Also, between August 2012–February 2013, he was a visiting student at the Telecommunications Technology Centre of Catalonia, Barcelona, Spain. In January 2014, he was selected as an Exemplary Reviewer of the IEEE Communication Letters Reviewer Appreciation Program 2013. His main research interests include signal processing and applied statistics with focus on the optimization and performance analysis of MIMO wireless communication systems.

Marco Di Renzo (S'05-AM'07-M'09-SM'14) received the Laurea (cum laude) and the Ph.D. degrees in Electrical and Information Engineering from the Department of Electrical and Information Engineering, University of L'Aquila, Italy, in April 2003 and in January 2007. In October 2013, he received the Habilitation à Diriger des Recherches (HDR) from the University of Paris-Sud XI, Paris, France.

Since January 2010, he has been a Tenured Academic Researcher ("Chargé de Recherche Titulaire") with the French National Center for Scientific Research (CNRS), as well as a faculty member of the Laboratory of Signals and Systems (L2S), a joint research laboratory of the CNRS, the École Supérieure d'Électricité (SUPÉLEC) and the University of Paris-Sud XI, Paris, France. His main research interests are in the area of wireless communications theory.

Dr. Di Renzo is a recipient of several awards, which include a special mention for the outstanding five-year (1997-2003) academic career, University of L'Aquila, Italy; the THALES Communications fellowship (2003-2006), University of L'Aquila, Italy; the 2004 Best Spin-Off Company Award, Abruzzo Region, Italy; the 2006 DEWS Travel Grant Award, University of L'Aquila, Italy; the 2008 Torres Quevedo Award, Ministry of Science and Innovation, Spain; the "Dérogation pour l'Encadrement de Thèse" (2010), University of Paris-Sud XI, France; the 2012 IEEE CAMAD Best Paper Award; the 2012 IEEE WIRELESS COMMUNICATIONS LETTERS Exemplary Reviewer Award; the 2013 IEEE VTC-Fall Best Student Paper Award; the 2013 Network of Excellence NEWCOM# Best Paper Award; the 2013 IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY Top Reviewer Award; the 2013 IEEE-COMSOC Best Young Researcher Award for Europe, Middle East and Africa (EMEA Region); and the 2014 IEEE ICNC Single Best Paper Award Nomination (Wireless Communications Symposium). Currently, he serves an an Editor of the IEEE COMMUNICATIONS LETTERS and of the IEEE TRANSACTIONS ON COMMUNICATIONS (Heterogeneous Networks Modeling and Analysis).

Harald Haas holds the Chair of Mobile Communications at the University of Edinburgh. His main research interests are in optical wireless communications, hybrid optical wireless and RF communications, spatial modulation and massive MIMO, and interference coordination in wireless networks. He first introduced and coined "spatial modulation" and "Li-Fi". Li-Fi was listed among the 50 best inventions in TIME Magazine 2011. Professor Haas was an invited speaker at TED Global 2011, and his talk has been watched online more than 1.5 million times. He is co-founder and chief scientific officer (CSO) of pureLiFi Ltd. Professor Haas holds 26 patents and has more than 20 pending patent applications. He has published 280 conference and journal papers including a paper in Science. He was co-recipient of a best paper award at the IEEE Vehicular Technology Conference in Las Vegas in 2013. In 2012, he was the only recipient of the prestigious Established Career Fellowship from the EPSRC (Engineering and Physical Sciences Research Council) within Information and Communications Technology in the UK. Professor Haas is recipient of the Tam Dalyell Prize 2013 awarded by the University of Edinburgh for excellence in engaging the public with science. In 2014, he was selected by EPSRC as one of ten RISE (Recognising Inspirational Scientists and Engineers) Leaders.

F.3 IEEE 75th Vehicular Technology Conference (VTC 2012-Spring)

Transmit Precoding for Receive Spatial Modulation Using Imperfect Channel Knowledge

Athanasios Stavridis*, Sinan Sinanović*, Marco Di Renzo†, and Harald Haas*

*Institute for Digital Communications
Joint Research Institute for Signal and Image Processing
School of Engineering
The University of Edinburgh
EH9 3JL, Edinburgh, UK
{a.stavridis, s.sinanovic, h.haas}@ed.ac.uk

†Laboratory of Signals and Systems (L2S),
French National Center for Scientific Research (CNRS)
École Supérieure d'Électricité (SUPÉLEC),
University of Paris–Sud XI (UPS)
3 rue Joliot–Curie, 91192 Gif–sur–Yvette (Paris), France
marco.direnzo@lss.supelec.fr

***Abstract*—In this paper, motivated by the relatively new concept of Spatial Modulation (SM), we are addressing the problem of Receive-Spatial Modulation (R-SM) under two cases of imperfect channel knowledge at the transmitter side. In the first case, we adopt a statistical model for the channel uncertainties, whereas in the second one a worst-case approach is followed and the channel uncertainties are expressed as a bounded set. Based on Zero-Forcing (ZF) precoding and using standard tools from optimization theory, we derive closed form solutions that turn out to be robust. Simulation results show that the proposed schemes have a performance close to the perfect channel knowledge scenario in low and mid-low SNRs. Furthermore, these designs can be applied to wide range of channels with different correlation states combined with a transmit power gain.**

I. Introduction

In the last years, we have been facing an increasing demand for wireless system throughput. This demand creates the need for spectrally efficient and reliable radio communication systems which will meet all the Quality of Services (QOS) standards combined with increased energy efficiency. The most promising way to fulfil this need is Multiple-Input Multiple-Output (MIMO) technology where multiple antennas are employed at the receiver and at the transmitter [1]. For example, the well known Vertical Bell Laboratories Layered Space-Time (V-BLAST) architecture achieves high data rates by using spatial multiplexing [2]. The main drawbacks that V-BLAST faces are the need for Inter-Antenna Synchronization (IAS) and Inter-Channel Interference (ICI) cancellation. As a result, the computational complexity of the receiver is high.

Recently, a new MIMO scheme has been proposed named Spatial Modulation (SM) [3], [4], [5]. SM avoids ICI by using only one antenna during transmission period to convey information, all the other antennas are not activated. With this method we not only avoid ICI but also gain a second mechanism of conveying information in addition to the classical M-ary Quadrature Amplitude Modulation (QAM) constellation. This second mechanism is the index of the transmitting antenna. Furthermore, SM does not require IAS. In addition, a special case of SM called Space Shift Keying (SSK) is proposed in [7]. Other SM-like schemes include [8] and [9], where both space and time dimension is used in order to form an extended concept of SSK. It is notable that SM has the capability to outperform other MIMO systems like V-BLAST [4]-[9].

In [10] a variation of SM is proposed called pre-processing aided spatial modulation. Using Channel State Information at the Transmitter side (CSIT) and precoding it is formed a Receive-Spatial Modulation (R-SM) scheme. Instead of activating only one antenna at the transmitter as SM, R-SM receives the transmitted signal at only one antenna at the receiver. Explicitly, the transmitter activates all of its antennas and using precoding targets at only one receive antenna. Thus, the index of the receiving antenna can be used by the transmitter as the additional mechanism of conveying information. In this work, we will use the name of R-SM for the pre-processing aided spatial modulation for reasons of simplicity.

Generally, precoding using CSIT on MIMO systems has been used either for receiver simplification and power consuming benefits for the case of downlink, or capacity gain. The key factor for precoding is the CSIT, which can be obtained either using low rate feedback from the receiver or using the reciprocity principle. In most cases, Channel Side Information at the Receiver (CSIR) is obtained using a training sequence and can be assumed to be accurate. In contrast, supplying the transmitter with accurate CSIT is a difficult task. Thus, the precoding designs should be robust under imperfect or partial CSIT. Generally, the design of robust precoders is divided in two categories. The first category models the channel uncertainties as Random Variables (RV) and solves the resulting optimization problem using a statistical approach [11]. In contrast, the second category makes the assumption that the channel uncertainty is a bounded set and a worst-case approach is followed [12].

In this work, we use Stochastic Robust Approximation (SRA) and Worst-Case Robust Approximation (WCRA) [15] in order to form a R-SM scheme similar to [10] that remains robust under imperfect CSIT. The motivation to base our design on SRA and WCRA is [13], where these optimization methods are used in MIMO precoding under frequency selec-

978-1-4673-0990-5/12/$31.00 ©2012 IEEE

tive channel and imperfect CSIT. Similar to [13], the resulting solutions have a closed form and appears to have competitive performance under different correlation scenarios combined with a transmit power gain. Finally, R-SM has a complexity gain at the receiver compared to SM, due to the simplification of the detector.

The rest of the paper is organized as follows: section II presents the system model and the problem statement. The proposed precoding schemes are presented in section III and their performances are shown in section IV. Finally, we conclude the paper in section V.

Notation: In the following, lowercase bold letters denote vectors and uppercase bolt letters denote matrices. $(\cdot)^T$, $(\cdot)^H$, tr$(\cdot)$ and $\mathbf{A}^{1/2}$ denotes transpose, Hermitian transpose, matrix trace and the square root of $\mathbf{A}$, respectively. E$[\cdot]$ is the mean value of a RV. Finally, a complex Gaussian distribution with mean m and variance σ_C^2 is represented as $CN(m, \sigma_C^2)$, where its real and imaginary part are independent and identically distributed (i.i.d.) Gaussian RV with distribution $N(m, \frac{\sigma_C^2}{2})$.

II. System Model and Problem Statement

A. System Model

A MIMO frequency flat system is considered with N_t transmit antennas and N_r receive antennas, with $N_t \geq N_r$. When a precoder is applied at the transmitter, the system equation at every symbol period is given as

$$\mathbf{y} = \mathbf{HPx} + \mathbf{w}, \tag{1}$$

where $\mathbf{H}$ is a $N_r \times N_t$ matrix representing the wireless MIMO channel, $\mathbf{P}$ is the precoding matrix of size $N_t \times N_r$, $\mathbf{x}$ is the $N_r \times 1$ transmitted signal vector and $\mathbf{w} \in \mathcal{C}^{N_r}$ is the zero mean vector of additive complex Gaussian noise with i.i.d. elements and variance of σ_w^2. A strategy for designing the precoding matrix $\mathbf{P}$, that aims at the elimination of Inter-Channel Interference (ICI), is the ZF method. In this case the precoding matrix is just the pseudo-inverse of the channel matrix $\mathbf{H}$. Using Singular Value Decomposition (SVD) the precoding matrix is equal to $\mathbf{P} = \mathbf{V\Sigma}^{-1}\mathbf{U}^H$ (where $\mathbf{H} = \mathbf{U\Sigma V}^H$).

In this paper, we use ZF in order to form a R-SM system as in [10]. This is done by using the precoding matrix $\mathbf{P}$ to eliminate the effect of MIMO channel and the appropriate formation of the transmitted symbol $\mathbf{x}$.

The system model we use is the same as in [10]. Let us assume that the transmitted vector $\mathbf{x}$ has the form of $\mathbf{x} = \mathbf{e}_i s_k$, where the vector $\mathbf{e}_i = [0, ..., 0, 1, 0, ...0]^T$, $i = 1, ..., N_r$, has length of N_r and all of its elements are zero except from the one at the i-th position which has the value of 1. $s_k \in \{s_1, ..., s_M\}$ represents the transmitted symbol selected from a constellation like M-QAM. Furthermore, if we assume perfect channel knowledge at the transmitter and the ZF precoding matrix as described before, the system equation has the following form:

$$\mathbf{y} = \underbrace{\mathbf{HP}}_{\mathbf{I}_{N_r, N_r}} \mathbf{e}_i s_k + \mathbf{w}. \tag{2}$$

In this scenario, the observation at each of the N_r receive antennas is given as

$$\begin{array}{ll} y_j = s_k + w_j, & j = i, \\ y_j = w_j, & j \neq i, \end{array} \tag{3}$$

$j = 1, ..., N_r$, which means that only one antenna receives the transmitted symbol s_k degraded from Gaussian noise and all the other receive antennas are facing only Gaussian noise.

If we configure our system so that the number of receive antennas is a power of 2, it is possible to map $k_1 = \log_2 N_r$ bits on the index of the i-th receive antenna while $k_2 = \log_2 M$ bits are mapped on the transmitted symbol s_k. In this way, we have formed the same R-SM system as in [10] which conveys $k = k_1 + k_2$ bits every symbol period.

Because of this, it is straightforward to use a ML detector (similar to the one used in [10])

$$(\hat{i}, \hat{s}_k) = \arg\min_{i, s_k} \|\mathbf{y} - \mathbf{e}_i s_k\|_2^2, \tag{4}$$

which minimizes the Euclidean norm over all possibles combinations of $i = 1, ..., N_r$ and $s_k \in \{s_1, ..., s_{k_2}\}$. As we can see from (4), the detector does not requires any channel knowledge in order to detect the received $(\hat{i}, \hat{s}_k)$.

B. Complexity of ML Detection

Using the zero structure of $\mathbf{e}_i$, the ML detector can be rewritten as

$$(\hat{i}, \hat{s}_k) = \arg\min_{i, s_k} |s_k|^2 - 2\text{Re}\{y_i^* s_k\}. \tag{5}$$

If we consider that $|s_k|^2$ is pre-computed and stored at the memory, the evaluation of (5) requires $N_r M$ complex multiplications, $N_r M$ real multiplications and $N_r M$ real additions.

Especially, for the case of Receive-Space Shift Keying (R-SSK), where $s_k = 1$ and only the index of the receive antenna is used as the mechanism of conveying information, the ML detector can be further simplified as

$$\hat{i} = \arg\max_i \text{Re}\{y_i\}, \tag{6}$$

which has the complexity of only N_r real comparisons.

Hence, if we compare the complexity of R-SM detection with the one of SM's optimal detector [6], which has the complexity of $2N_t N_r M - 2N_t M$ complex additions and $2N_t N_r M + N_t M$ complex multiplications plus the complexity of $N_t M$ real additions and $N_t M$ real multiplications, we can see a notable complexity simplification.

C. Objective Function

In this subsection, we present the objective function that we use in the next section to design the precoding matrix $\mathbf{P}$, using partial channel knowledge.

In order to design the ZF precoding matrix of (1) we can compute the pseudo-inverse matrix of $\mathbf{H}$ using SVD. An alternative way is to solve the following N_r minimization problems

$$\min_{\mathbf{p}_l} \|\mathbf{H}\mathbf{p}_l - \mathbf{e}_l\|_2^2, \ \forall l = 1, ..., N_r, \tag{7}$$

where $\mathbf{p}_l$ is the l-th column of the precoding matrix $\mathbf{P} = [\mathbf{p}_1, ..., \mathbf{p}_{N_r}]$ and $\mathbf{e}_l$ is the l-th column of the identity matrix $\mathbf{I}_{N_r,N_r} = [\mathbf{e}_1, ..., \mathbf{e}_{N_r}]$. When perfect channel knowledge is available at the transmitter, the solution of the previous minimization problems is trivial. Our target in the next section is to solve the minimization problems of (7) using imperfect channel knowledge.

III. Precoding Using Imperfect Channel Knowledge

A. Imperfect Channel Knowledge at the Transmitter

In realistic scenarios providing the precoder with perfect channel knowledge is almost impossible. A common assumption is to model MIMO channel as $\mathbf{H} = \overline{\mathbf{H}} + \widetilde{\mathbf{H}}$ [14]. In this model, $\overline{\mathbf{H}}$ represents the long term channel evolution which can be accurately acquired by the transmitter. Whereas, $\widetilde{\mathbf{H}}$ denotes the channel for which only some kind of statistical or worst-case knowledge is considered possible. Usually $\widetilde{\mathbf{H}}$ is modeled as Zero Mean Circular Symmetric Complex Gaussian Matrix (ZMCSCG) with i.i.d. elements of variance $\sigma^2_{\widetilde{\mathbf{H}}}$.

In the following two subsections, the minimization problems of (7) are solved using two methods from optimization theory, SRA and WCRA [15]. SRA uses a statistical approach and solves the minimization problems over their expectation. The partial knowledge that SRA employs is the matrix $\overline{\mathbf{H}}$ and the matrix $\mathrm{E}[\widetilde{\mathbf{H}}^{\mathrm{H}}\widetilde{\mathbf{H}}]$. On the other hand, WCRA is a worst-case approach and solves the same minimization problems assuming full knowledge of $\overline{\mathbf{H}}$ and α which is defined as $\|\widetilde{\mathbf{H}}\|_2 \leq \alpha$.

Both designs are using the long term channel evolution $\overline{\mathbf{H}}$ which is a slow varying characteristic. Furthermore, SRA uses $\mathrm{E}[\widetilde{\mathbf{H}}^{\mathrm{H}}\widetilde{\mathbf{H}}]$ which is also a slow varying statistic [16] and in some cases, as it is described in the next subsection, has a structured form. On the other hand, WCRA uses the additional information of α which is just a scalar. Thus, our schemes enjoy the merit of reduced feedback from the receiver.

B. ZF-like Precoding Based on SRA

SRA solves the minimization problems of (7) over its expectation [15],

$$\min_{\mathbf{p}_l} \mathrm{E}\left[\|\mathbf{H}\mathbf{p}_l - \mathbf{e}_l\|_2^2\right], \ \forall l = 1, ..., N_r. \tag{8}$$

If we use the fact that the MIMO channel can be written as $\mathbf{H} = \overline{\mathbf{H}} + \widetilde{\mathbf{H}}$, where $\mathrm{E}[\widetilde{\mathbf{H}}] = \mathbf{0}$, (8) can be reformulated as

$$\min_{\mathbf{p}_l} \left\{ \|\overline{\mathbf{H}}\mathbf{p}_l - \mathbf{e}_l\|_2^2 + \mathbf{p}_l^H \mathrm{E}\left[\widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}}\right]\mathbf{p}_l \right\}, \tag{9}$$

which is obviously a convex optimization problem (sum of quadratic functions) and can be solved by setting its gradient equal to zero. The analytical solution is formed as

$$\mathbf{p}_l = \left[\overline{\mathbf{H}}^H\overline{\mathbf{H}} + \mathrm{E}\left[\widetilde{\mathbf{H}}^{\mathrm{H}}\widetilde{\mathbf{H}}\right]\right]^{-1}\overline{\mathbf{H}}^H\mathbf{e}_l. \tag{10}$$

Inspecting (10), we can see that the computation of $\mathbf{p}_l$ requires the knowledge of $\overline{\mathbf{H}}$ and the matrix $\mathrm{E}[\widetilde{\mathbf{H}}^{\mathrm{H}}\widetilde{\mathbf{H}}]$ at the transmitter. Finally, it is worth mentioning that in many scenarios $\mathrm{E}[\widetilde{\mathbf{H}}^{\mathrm{H}}\widetilde{\mathbf{H}}]$ has a structured form, which can simplify the hardware design and reduce the feedback from the receiver. For example, in the case of the uncorrelated channel $\mathbf{H}$, with $\tilde{h}_{i,j} \in CN(0, \sigma_{\widetilde{\mathbf{H}}})$, $\mathrm{E}[\widetilde{\mathbf{H}}^{\mathrm{H}}\widetilde{\mathbf{H}}]$ is the identity matrix multiplied by $N_r\sigma^2_{\widetilde{\mathbf{H}}}$.

C. ZF-like Precoding Based on WCRA

In this subsection we use WCRA [15] in order to design the precoding matrix $\mathbf{P}$. We define the non-empty and bounded set $\Phi \subseteq \mathbf{C}^{N_r,N_t}$ which represents all the possible values of the channel matrix $\mathbf{H}$. Given a feasible precoding vector $\mathbf{p}_l$ the worst case error can be formulated as $e_{\mathrm{wc}}(\mathbf{p}_l) = \sup\left[\|\mathbf{H}\mathbf{p}_l - \mathbf{e}_l\|_2 | \mathbf{H} \in \Phi\right]$. Our target here is to design a precoder which minimizes the $e_{\mathrm{wc}}(\mathbf{p}_l)$. In this case, the minimization problem can be formulated as

$$\min_{\mathbf{p}_l} \sup\left[\|\mathbf{H}\mathbf{p}_l - \mathbf{e}_l\|_2 | \mathbf{H} \in \Phi\right] \ \forall l = 1, ..., N_r. \tag{11}$$

One possible solution to the minimization problem of (11) can be reached using the Norm Bound Error (NBE) [15]. In this method, the uncertainty of $\widetilde{\mathbf{H}}$ is considered within a norm ball of radius α and the set Φ is written as $\Phi = \left\{\mathbf{H} = \overline{\mathbf{H}} + \widetilde{\mathbf{H}} | \|\widetilde{\mathbf{H}}\|_2 \leq \alpha\right\}$, where $\alpha > 0$. Let $e^{\mathrm{NBE}}_{\mathrm{wc}}(\mathbf{p}_l) = \sup\{\|\overline{\mathbf{H}}\mathbf{p}_l - \mathbf{e}_l + \widetilde{\mathbf{H}}\mathbf{p}_l\|_2 | \ \|\widetilde{\mathbf{H}}\|_2 \leq \alpha\}$ be the worst-case error given the precoding vector $\mathbf{p}_l$. It is easily to show [15] that $e^{\mathrm{NBE}}_{\mathrm{wc}}(\mathbf{p}_l)$ is equal to $e^{\mathrm{NBE}}_{\mathrm{wc}}(\mathbf{p}_l) = \|\overline{\mathbf{H}}\mathbf{p}_l - \mathbf{e}_l\|_2 + \alpha\|\mathbf{p}_l\|_2$ and it is attained for $\widetilde{\mathbf{H}} = \alpha\mathbf{u}\mathbf{v}^H$ where $\mathbf{u} = \frac{\overline{\mathbf{H}}\mathbf{p}_l - \mathbf{e}_l}{\|\overline{\mathbf{H}}\mathbf{p}_l - \mathbf{e}_l\|_2}$ and $\mathbf{v} = \frac{\mathbf{p}_l}{\|\mathbf{p}_l\|_2}$, given that $\overline{\mathbf{H}}\mathbf{p}_l - \mathbf{e}_l \neq 0$ and $\mathbf{p}_l \neq \mathbf{0}$ [15]. Thus the minimization problem of (11) can be reformulated as $\min_{\mathbf{p}_l} \|\overline{\mathbf{H}}\mathbf{p}_l - \mathbf{e}_l\|_2 + \alpha\|\mathbf{p}_l\|_2$, which can be re-written in a Tikhonov Regularization form for some value of β [15]

$$\min_{\mathbf{p}_l} \|\overline{\mathbf{H}}\mathbf{p}_l - \mathbf{e}_l\|_2^2 + \beta\|\mathbf{p}_l\|_2^2. \tag{12}$$

Again, the minimization problem of (12) is convex because $e^{\mathrm{NBE}}_{\mathrm{wc}}(\mathbf{p}_l)$ is the sum of quadratic functions. Thus the solution can be reached using the gradient condition

$$\mathbf{p}_l = \left[\overline{\mathbf{H}}^H\overline{\mathbf{H}} + \beta\mathbf{I}\right]^{-1}\overline{\mathbf{H}}^H\mathbf{e}_l \ \forall l = 1, ..., N_r. \tag{13}$$

D. Precoding in the Presence of Transmit and Receive Space Correlations

It is clear that both problems of (9) and (12) are different forms of Tikhonov Regularization. A valuable property of Tikhonov Regularization theory is that it does not pose any rank restriction on the involved matrices $\overline{\mathbf{H}}$ and $\widetilde{\mathbf{H}}$ as long as the matrices $\overline{\mathbf{H}}^H\overline{\mathbf{H}} + \mathrm{E}\left[\widetilde{\mathbf{H}}^{\mathrm{H}}\widetilde{\mathbf{H}}\right]$ and $\overline{\mathbf{H}}^H\overline{\mathbf{H}} + \beta\mathbf{I}$ are positive definite [15]. Thus, our analytical solutions of (10) and (13) enjoy the additional merit of being applicable to spatially correlated channels.

In this paper we employ the Kronecker correlation model [17]. Under this correlation model the MIMO channel can be rewritten as

$$\mathbf{H} = \mathbf{R}_R^{1/2}\mathbf{H}_w(\mathbf{R}_T^{1/2})^T, \tag{14}$$

where $\mathbf{H}_w$ is a ZMCSCG matrix with i.i.d elements and variance of $\sigma^2_{\mathbf{H}_w} = 1$. The matrices $\mathbf{R}_T$ and $\mathbf{R}_R$ represent the transmit spatial correlation matrix and the receive spatial correlation matrix, respectively. Usually, the entries of the spatial correlation matrices $\mathbf{R}_R$ and $\mathbf{R}_T$ are generated using an exponential model with $R_T(i,j) = \rho_t^{|i-j|}$ and

$R_R(i,j) = \rho_r^{|i-j|}$, where $0 \leq \rho_t, \rho_r \leq 1$. Values of ρ_t and ρ_r close to 0 mean low correlation, whereas values closely to 1 mean high correlations.

If we combine the Kronecker correlation model of (14) with the model of partial channel knowledge described at subsection III-A the MIMO channel can be written as

$$\mathbf{H} = \mathbf{R}_R^{1/2}\left(\overline{\mathbf{H}} + \widetilde{\mathbf{H}}\right)(\mathbf{R}_T^{1/2})^T,$$

where the matrix $\mathbf{R}_R^{1/2}\overline{\mathbf{H}}(\mathbf{R}_T^{1/2})^T$ represents the full known part of the channel and the matrix $\mathbf{R}_R^{1/2}\widetilde{\mathbf{H}}(\mathbf{R}_T^{1/2})^T$ represents the partial known part of the channel.

Inspecting (10) and (13) we can see that only the analytical form of the design based on SRA is affected by the correlated channel. This is because SRA requires the computation of $\mathrm{E}[\widetilde{\mathbf{H}}^{\mathrm{H}}\widetilde{\mathbf{H}}]$ which can be shown that it takes the form of $\mathrm{E}[\widetilde{\mathbf{H}}^{\mathrm{H}}\widetilde{\mathbf{H}}] = \sigma_{\widetilde{\mathbf{H}}}^2 \mathrm{tr}(\mathbf{R}_{\mathrm{R}})\mathbf{R}_{\mathrm{T}}^{1/2}(\mathbf{R}_{\mathrm{T}}^{1/2})^{\mathrm{T}}$.

E. Transmit Power Analysis

ZF precoding has the disadvantage of increasing the transmit power significantly in order to overcome deep fades. In real systems, this is something that should be avoided. A common solution is to impose a transmit power constraint $\|\mathbf{P}\|_F^2 = \sum_{l=1}^{N_r} \|\mathbf{p}_l\|_2^2 \leq p_t$, where $\|\cdot\|_F^2$ is the Frobenius norm and p_t is the transmit power constraint.

SRA and WCRA are forms of Tikhonov Regularization problem. Tikhonov Regularization theory is the most common form of regularization that penalize big values of $\|\mathbf{p}_l\|_2^2$ [15]. As a consequence, our precoding schemes achieve solutions with small values of $\|\mathbf{p}_l\|_2^2$ depending on the value of channel uncertainty $\sigma_{\widetilde{\mathbf{H}}}^2$. In SRA and WCRA, increased values of $\sigma_{\widetilde{\mathbf{H}}}^2$ produce precoding vectors $\mathbf{p}_l$ with small Euclidean norm [15]. Hence, our systems indirectly satisfy the previous transmit power constraint . In the next section, we study the behavior of the transmitted power of our schemes using numerical simulations.

IV. Simulation Results

In this section, we present simulation results that demonstrate the performance of the proposed precoding methods. The system configuration assumes $N_t = 4$ transmit antennas and $N_r = 2$ receive antennas. Furthermore, every transmitted symbol $s_k \in \{s_1, ..., s_M\}$ is selected from a 4-QAM constellation with average transmitted power equal to 1. As a consequence, every symbol period a total of $k_1 + k_2 = 3$ bits are transmitted. The MIMO channel $\mathbf{H}$ is generated as the sum of two ZMCSCG matrices ($\overline{\mathbf{H}}$ and $\widetilde{\mathbf{H}}$) whose elements follow i.i.d. $\overline{h}_{i,j} \sim CN(0, 0.99)$ and $\tilde{h}_{i,j} \sim CN(0, 0.01)$ distribution such that $h_{i,j} \sim CN(0,1)$. The entries of $\overline{\mathbf{H}}$ are refreshed every 100 symbol periods whereas the entries of $\widetilde{\mathbf{H}}$ are refreshed every symbol period. In addition, the entries of the spatial correlation matrices $\mathbf{R}_R$ and $\mathbf{R}_T$ are generated using an exponential model with $\rho_t = \rho_r = \rho$, where ρ takes values from the set $\{0.1, 0.3, 0.5, 0.9\}$. Finally, Signal-to-Noise Ratio (SNR) per bit is defined as $\mathrm{SNR}_{\mathrm{bit,dB}} = 10\log\frac{\mathrm{tr}(\mathbf{P}^H\mathbf{P})\sigma_s^2/[N_r(k_1+k_2)]}{\sigma_w^2}$, where the variance of the transmitted symbol s_k is equal to $\sigma_s^2 = 1$[1].

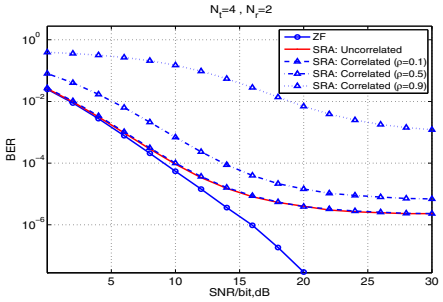


Fig. 1. BER versus SNR for SRA precoder for $\sigma_{\widetilde{\mathbf{H}}}^2 = 0.01$

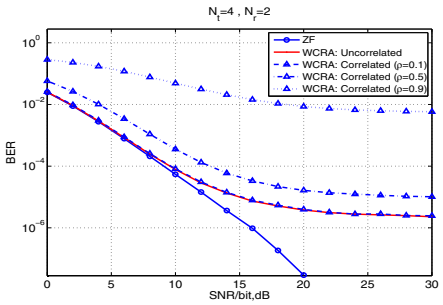


Fig. 2. BER versus SNR for WCRA precoder for $\sigma_{\widetilde{\mathbf{H}}}^2 = 0.01$

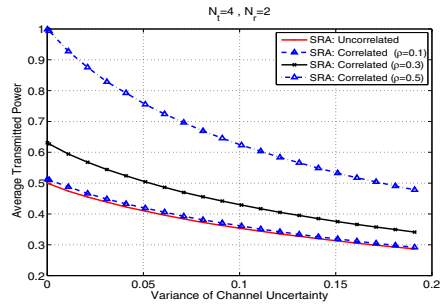


Fig. 3. Average transmitted power versus $\sigma_{\widetilde{\mathbf{H}}}^2$ for SRA precoder

The BER performance of the two proposed schemes is presented in Fig. 1 and Fig. 2 for the cases of uncorrelated and correlated channel. Especially, for the correlated case we assume spatial correlated channels with low correlation ($\rho = 0.1$), with intermediate correlation ($\rho = 0.5$) and with

[1]The transmitted signal power is divided by N_t because every symbol period only one column of the precoder $\mathbf{P}$ is activated.

high correlation ($\rho = 0.9$). In addition, in the same figures we present the BER performance of the ZF precoder under full channel knowledge. In order to have a fair comparison between the proposed designs and the ZF precoder, the channel that the ZF precoder faces has the same statistical characteristics as our schemes.

Inspecting Fig. 1 and Fig. 2, we can see that under the uncorrelated channel the proposed designs have similar performance to the ZF precoder at the low and mid-low SNRs. Clearly, at the mid-high and high SNRs ZF precoder outperforms our scheme. In these SNRs the main degradation factor for our schemes is the residual ICI, ZF precoder has better performance due to accurate channel knowledge. Furthermore, it seems that our systems, due to the residual ICI, face a bottleneck at 20 dB, where higher values of SNR do not affect their performance rapidly. The way that correlation affects the performance of our schemes can be divided in two cases. In the low correlation case ($\rho = 0.1$), we can see that there is no performance difference between the uncorrelated and correlated channel. In the second case, as correlation increases the performance of our systems deteriorates. For the intermediate level of correlation ($\rho = 0.5$) there is a need of 10 dB in order to achieve a BER performance of practical interest. For the high level of correlation ($\rho = 0.9$) this need is increased to 22 dB.

Fig. 3 and Fig. 4 present the average transmitted power versus the variance of channel uncertainty for different correlation scenarios. As it can be seen from these figures, as the channel uncertainty $\sigma^2_{\tilde{\mathbf{H}}}$ increases the transmitted power decreases. This means that when the channel knowledge at the transmitter becomes more inaccurate, the transmitter reduces the transmitted power in order to avoid further degradation. Although increasing channel correlation requires higher average transmit power, the proposed new scheme requires for low ($\rho = 0.1$) correlation almost the same power as the uncorrelated case. Finally, the results offer another indirect comparison between our precoding designs and the ZF precoder using perfect channel knowledge. When $\sigma^2_{\tilde{\mathbf{H}}} \to 0$, SRA and WCRA reduce to ZF precoder with full channel knowledge. Clearly, there is a significant power gain for our schemes compared to ZF precoder especially for higher values of $\sigma^2_{\tilde{\mathbf{H}}}$.

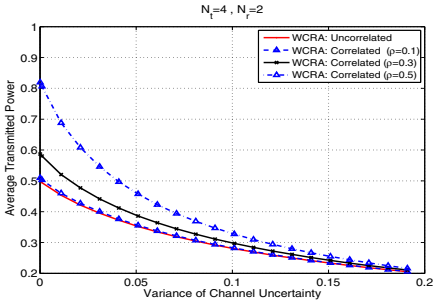


Fig. 4. Average transmitted power versus $\sigma^2_{\tilde{\mathbf{H}}}$ for WCRA precoder

V. Conclusions

This paper treats the problem of R-SM under imperfect CSIT. Using two different forms of imperfect channel knowledge we formulate two precoding schemes for R-SM. By expressing the resultant optimization problems into Tikhonov regularization forms we achieve a transmit power gain and the capability to treat correlated channels. In this way, we propose two R-SM schemes that make use of practical forms of CSIT and have low detection complexity at the receiver. Our systems turns to have a well-behaved BER performance under different channel correlation states combined with a transmit power gain. Finally, the simulation results confirm the advantages of the new schemes.

Acknowledgement

This work has been funded by the Research Project GREENET (PITN-GA-2010-264759).

References

[1] E. Telatar, *Capacity of the Multiple Antenna Gaussian Channel*, European Transactions on Telecommunications, 10(6), pp. 585-595, 1999.

[2] P. Wolniansky, G. Foschini, G. Golden, and R. Valenzuela, *V-BLAST: An Architecture for Realizing Very High Data Rates Over the Rich-Scattering Wireless Channel*, in Proc. International Symp. Signals, Systems, Electronics (ISSSE98), Pisa, Italy, pp. 295-300, September 1998.

[3] M.D. Renzo, H. Haas, P.M. Grant, *Spatial Modulation for Multiple-Antenna Wireless Systems: A Survey*, IEEE Communications Magazine, vol. 49, pp. 182 - 191, 2011.

[4] R. Mesleh, H. Haas, C. Ahn, and S. Yun, *Spatial modulation - A New Low Complexity Spectral Efficiency Enhancing Technique*, in Proc. First International Conf. Commun. Netw., Beijing, China, October 2006, pp. 1-5.

[5] R. Mesleh, H. Haas, S. Sinanovic, C. Ahn, S. Yun, *Spatial Modulation*, IEEE Transactions on Vehicular Technology, vol. 57, no. 4, pp. 2228-2242, 2008.

[6] J. Jeganathan, A. Ghrayeb, L. Szczecinski, *Spatial Modulation: Optimal Detection and Performance Analysis*, IEEE Commun. Letters, vol.12, no.8, pp. 545-547, August 2008.

[7] J. Jeganathan, A. Ghrayeb, L. Szczecinski, A. Ceron, *Space Shift Keying Modulation for MIMO Channels*, IEEE Trans. on Wireless Commun., vol.8, no.7, pp. 3692- 3703, July 2009.

[8] S. Sugiura, S. Chen, L. Hanzo, *Coherent and Differential Space-Time Shift Keying: A Dispersion Matrix Approach*, IEEE Trans. on Commun., vol.58, no.11, pp. 3219-3230, November 2010

[9] S. Sugiura, S. Chen, H. Haas, P.M. Grant, L. Hanzo, *Coherent Versus Non-Coherent Decode-and-Forward Relaying Aided Cooperative Space-Time Shift Keying*, IEEE Trans. on Commun., vol. 59, no. 6, pp. 1707-1719, June 2011.

[10] L.-L. Yang, *Transmitter Preprocessing Aided Spatial Modulation for Multiple-Input Multiple-Output Systems*, IEEE 73th Vehicular Technology Conference: VTC2011-Spring, 15-18 May, 2011, Budapest, Hungary.

[11] X. Zhang, D. P. Palomar and B. Ottersten, *Statistically Robust Design of Linear MIMO Transceivers*, IEEE Trans. on Signal Processing, vol. 56, no. 8, pp. 3678–3689, August 2008.

[12] Y. Guo, B. C. Levy, *Worst-Case MSE Precoder Design for Imperfectly Known MIMO Communications Channels*, IEEE Trans. on Signal Processing, vol. 53, no. 8, pp. 2918-2930, 2005.

[13] A. Stavridis, S. Karachontzitis, K. Berberidis, *Bezout-Based Robust Precoding for MIMO Frequency Selective Channel Using Imperfect Channel Knowledge*, in Proc. EUSIPCO, Barcelona, 2011.

[14] M. Vu, A. Paulraj, *A Robust Transmit CSI Framework with Applications in MIMO Wireless Precoding*, in Proc. 39th Asilomar Conf. on Signals, Systems, and Computers, pp. 623-627, November 2005.

[15] S. Boyd, L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[16] A. Paulraj, R. Nabar, D. Gore, *Introduction to Space-Time Wireless Communications*, Cambridge University Press, 2003.

[17] J. Kermoal, L. Schumacher, K.I. Pedersen, P. Mogensen, F. Frederiksen, *A Stochastic MIMO Radio Channel Model With Experimental Validation*, IEEE Journal on Selected Areas Communications, vol 20, pp. 1211-1226, 2002.

F.4 IEEE 17th Int. Workshop on Computer Aided Modeling and Design of Communication Links and Network (CAMAD)

2012 IEEE 17th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)

An Energy Saving Base Station Employing Spatial Modulation

Athanasios Stavridis*, Sinan Sinanovic*, Marco Di Renzo†, Harald Haas*, and Peter Grant*

*Institute for Digital Communications
Joint Research Institute for Signal and Image Processing
School of Engineering
The University of Edinburgh
{a.stavridis, s.sinanovic, h.haas, peter.grant}@ed.ac.uk

†Laboratory of Signals and Systems (L2S),
French National Center for Scientific Research (CNRS)
École Supérieure d'Électricité (SUPÉLEC),
University of Paris–Sud XI (UPS)
marco.direnzo@lss.supelec.fr

***Abstract*—In this paper, we evaluate the energy efficiency of a multi-antenna Base Station (BS) employing Spatial Modulation (SM). Taking advantage of the single Radio Frequency (RF) chain configuration of SM, we show that SM offers a significant total power reduction compared to other multi-RF chain Multiple-Input Multiple-Output (MIMO) architectures. For the same RF transmit power, we demonstrate that the total power saving of SM scales linearly with the number of RF chains required by other MIMO schemes. Furthermore, we clarify that for the same ergodic capacity, SM has a significant advantage in energy efficiency compared to the studied multi-RF chain MIMO configurations. In addition, for a number of transmit antennas larger than two, we establish that SM results in higher ergodic capacity than Space-Time Block-Coding (STBC), combined with significant power saving. For a BS with 8 transmit antenna, the achieved power saving of SM can reach up to almost 90%. Finally, based on simulation results, we demonstrate that the energy efficiency (bits/J) of the studied MIMO schemes exhibits a maximum as function of ergodic capacity, and it is further shown that this maximum is several times higher in SM compared to the studied state-of-the-art MIMO schemes.**

I. Introduction

The ever increasing need for wireless system throughput implies that future wireless networks must be spectrally efficient in order to meet all the Quality of Service (QoS) of data rate demanding applications like real time video streaming. Over the last fifteen years, research on MIMO wireless systems has shown that multi-antenna elements, both at the transmitter and the receiver, are the most promising solution for the next generation networks [1].

Unfortunately, the employment of multi-antenna elements comes with one major disadvantage, the requirement for multiple RF chains at both the transmitter and receiver. Firstly, RF chains are circuits that do not follow Moore's law in progressive improvement [2]. Hence, RF chains may increase significantly the cost of real system implementation. Secondly, RF chains contain Power Amplifiers (PAs) which are responsible for the 50-80% of the total power consumption in the transmitter [3]. Thus, increased number of PAs results in a decrease in energy efficiency. Therefore, there is a demand for MIMO schemes which employ a single RF chain. SM is an example that offers this property.

SM is a relatively new MIMO concept that enables high data rates while retaining the single RF chain configuration at the transmitter [4]. By splitting the transmitted bit sequence into two separate parts, SM achieves a multiplexing gain. The first part is encoded in the index of the active transmit antenna, while the second part is encoded via a conventional signal constellation point. The activation of only one transmit antenna totally avoids Inter-Channel Interference (ICI) at the receiver, resulting in a low complexity Maximum Likelihood (ML) detector. An additional merit of SM is that it does not require Inter-Antenna Synchronization (IAS) at the transmitter.

Until recently, research on energy efficiency was mainly focused on reducing and quantifying the RF transmit power. In this way, only a portion of the total transmit power is taken into account, which may give rise to misleading conclusions. A simple, yet insightful power model is presented in [5]. By taking into account that the power supply increases with the RF transmission power while there is also a constant input power which is independent of the RF transmit power a linear power model is derived and verified with real data. This model quantifies the total power consumption at the transmitter as opposed to only the RF transmit power.

In this paper, motivated by the potential of SM [4], we employ the power model from [5] in order to study the overall power performance of a BS employing SM. We show that SM requires less total power than any other multi-RF chain architecture, in order to transmit the same amount of RF power. Furthermore, we show that the capacity achieved by SM requires significantly less power than any other competitive MIMO, such as [6, 7]. Especially, for a BS with a number of transmit antennas more than two, SM is much more spectrally and power efficient than Space-Time Block-Coding (STBC). Finally, we study the energy efficiency (bits/J) of the previous schemes, and using simulation results show that SM is the most energy efficient among them.

The reminder of this paper is organized as follows: Section II presents SM and its information theoretic analysis is demonstrated in Section III. The employed power model is illustrated in Section IV. The evaluation of transmit power with respect to the total power consumption is presented in

978-1-4673-3125-8/12/$31.00 ©2012 IEEE

Section V. Moreover, Section VI compares SM with other MIMO architectures in terms of ergodic capacity versus total power consumption. The energy efficiency analysis of SM is demonstrated in Section VII. Finally, we conclude in Section VIII.

Notation: In the following, lowercase bold letters denote vectors and uppercase bold letters denote matrices. Furthermore, a complex Gaussian distribution with mean m and variance σ_C^2 is represented as $CN(m, \sigma_C^2)$, where its real and imaginary parts are independent and identically distributed (i.i.d.) Gaussian RV with distribution $N(m, \frac{\sigma_C^2}{2})$. Finally, $\|\cdot\|_2$ stands for the Euclidean norm.

II. Spatial Modulation

Our main interest in this paper is to analyze the total power consumption of a multi-antenna BS incorporating SM. We begin with a brief description of SM for the case of single user point-to-point communication [4]. We assume a transmitter with N_t antennas and a receiver with N_r antennas. In SM, the system equation is written as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}, \tag{1}$$

where $\mathbf{y}$ is the $N_r \times 1$ received signal vector, $\mathbf{H}$ is the $N_r \times N_t$ flat fading channel matrix, $\mathbf{x}$ is the $N_t \times 1$ transmitted signal vector, and $\mathbf{w} \in \mathcal{C}^{N_r}$ is the i.i.d. additive complex Gaussian noise with $w_i \sim CN(0, N)$. In addition, the transmitted vector is formulated as $\mathbf{x} = \mathbf{e}_i s_k$, where the vector $\mathbf{e}_i = [0, ..., 0, 1, 0, ..., 0]^T$, $i = 1, ..., N_t$, has length of N_t and all of its elements are zero except the one at the i-th position which has the value of 1. The position of the non-zero element of $\mathbf{e}_i$ corresponds to the position of the single activated antenna. Moreover, $s_k \in \{s_1, ..., s_M\}$ is a point selected from a signal constellation such as M-ary Quadrature Amplitude Modulation (QAM). Using this transmit signal formulation, a total of $k = \log_2 N_t + \log_2 M$ bits can be transmitted per symbol period. The first $k_1 = \log_2 N_t$ bits are allocated on the selection of the activated antenna and the last $k_2 = \log_2 M$ bits are allocated on the selection of the conventional constellation point.

At the receiver side, SM employs a low complexity optimal ML detector

$$(\hat{i}, \hat{s}_k) = \arg\min_{i, s_k} \|\mathbf{y} - \mathbf{H}\mathbf{e}_i s_k\|_2^2 \tag{2}$$

proposed in [8].

III. Capacity of Spatial Modulation

As it is described in Section II, SM conveys information using a two-dimensional space constituted from the constellation point s_k and the spatial domain in the form of the index of the activated transmit antenna. The information theoretic analysis of SM is presented in [9], where the capacity of SM is expressed as

$$C_{\text{total}} = W(C_1 + C_2), \tag{3}$$

where W is the transmission bandwidth, C_1 represents the capacity of the signal domain and C_2 denotes the capacity included in the spatial domain. The capacity for C_1 is given as

$$C_1 = \frac{1}{N_t} \sum_{i=1}^{N_t} \log_2 \left(1 + \frac{P\|\mathbf{h}_i\|_2^2}{N}\right), \tag{4}$$

where P represents the average transmitted power, $\mathbf{h}_i$ is the i-th column of the channel matrix, and N is the noise power. Moreover, given that the transmitted signal is selected as an i.i.d. complex Gaussian random variable, C_2 is obtained at the same paper as

$$C_2 = \frac{1}{N_t} \sum_{i=1}^{N_t} \int_y p(y|\mathbf{h}_i) \log_2 \frac{p(y|\mathbf{h}_i)}{p(y)} dy, \tag{5}$$

where

$$p(y|\mathbf{h}_i) = \frac{1}{\pi(P\|\mathbf{h}_i\|_2^2 + N)} \exp\left(-\frac{|y|^2}{P\|\mathbf{h}_i\|_2^2 + N}\right) \tag{6}$$

represents the probability distribution function (pdf) of the received signal with respect to all receive antennas. In addition,

$$p(y) = \frac{1}{N_t} \sum_{i=1}^{Nt} p(y|\mathbf{h}_i) \tag{7}$$

denotes the pdf of the average received signal.

Finally, we should note that [9] shows that SM achieves higher capacity than STBC for a number of transmit antennas greater than two ($N_t > 2$).

IV. Power Consumption Model

A modern multi-antenna BS consists of multiple transceivers. Each transceiver contains an Antenna Interface (AI), a PA, a Baseband Interface (BI), a DC-DC power supply, an AC-DC supply and a cooling system. Higher traffic load increases the power consumption on some of these components, whereas others keep their power consumption constant. The total power consumption of this type of BS is load dependent mainly because of the PAs.

A quite simple, yet very elegant power model that captures the traffic load power consumption behavior of a BS is proposed in [5]. In this model, the total power consumption is approximated using a linear function

$$P_{\text{supply}} = \begin{cases} N_{\text{RF}} P_0 + m N_{\text{RF}} P_{\text{t}}, & 0 < P_{\text{t}} \leq P_{\text{max}} \\ P_{\text{sleep}}, & P_{\text{t}} = 0, \end{cases} \tag{8}$$

where P_{supply} denotes the power supply, N_{RF} is the number of RF chains employed at the BS, P_0 is the power consumption at the minimum possible transmitted power, m is the slope that quantifies the load dependent power consumption, P_{t} is the transmitted power during transmission, P_{max} is the maximum allowed RF transmitted power, and P_{sleep} denotes the power supply when there is no transmission.

It is clear from (8) that the number of RF chains employed on a BS has a significant effect on the total power consumption. Increased value of N_{RF} is translated into increased power consumption.

Table I presents the power model parameters for different State-Of-The-Art (SOTA) BSs at the year of 2010 [5], [10].

TABLE I
POWER MODEL PARAMETERS FOR DIFFERENT BS (SOTA 2010) [5].

BS type	P_0 (W)	m	P_{max} (W)	P_{sleep} (W)
Macro	118.7	2.66	40.0	63
Micro	53.0	3.1	6.3	-
Pico	6.8	4.0	0.13	-
Femto	4.8	7.5	0.05	-

From now on, we term the values of Table I as the SOTA 2010 model. A prediction for the values of the power model of (8), at the year 2014, is also presented in [3]. We restrict our study to the SOTA 2010 model, because we wish to emphasize the total power consumption gain at the BS that can be achieved with today's hardware.

V. TOTAL BS POWER CONSUMPTION FOR DIFFERENT MIMO CONFIGURATIONS

In the following, we use the previous power model (Section IV) in order to evaluate the total power consumption of a BS with N_t antennas, assuming SM or any other Conventional MIMO (C-MIMO) scheme like Vertical-Bell Laboratories Layered Space-Time (V-BLAST) [11] or STBC [7]. Unfortunately, C-MIMO architectures require multiple RF chains at the transmitter (Fig. 2).

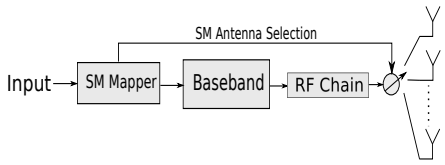


Fig. 1. Basic block diagram of a SM transmitter.

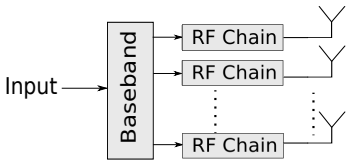


Fig. 2. Basic block diagram of a conventional MIMO transmitter.

SM is a single RF chain scheme (Fig. 1), and although it employs multiple antennas at the transmitter, the number of the utilized RF chains is $N_{\text{RF}} = 1$. Hence, (8) must be reformulated as

$$P_{\text{supply}}^{\text{SM}} = P_0 + mP_{\text{t}}. \tag{9}$$

In contrast, for the case of every other C-MIMO scheme where $N_t = N_{\text{RF}}$, (8) is written as

$$P_{\text{supply}}^{\text{CM}} = N_{\text{RF}}P_0 + m\underbrace{N_{\text{RF}}P_{\text{t}}}_{P_{\text{t}}^{\text{Total}}}, \tag{10}$$

where $P_{\text{t}}^{\text{Total}} = N_{\text{RF}}P_{\text{t}}$ is total power transmission from all antennas and P_{t} is the power transmission form a single antenna.

In order to evaluate the total power consumption of SM at a BS compared to C-MIMO scheme, we study two different scenarios. The first scenario assumes that the total power transmission $P_{\text{t}}^{\text{Total}}$ of a C-MIMO BS is equal to the power transmission of SM ($P_{\text{t}}^{\text{Total}} = P_{\text{t}}$). Using equations (9) and (10), it can be shown that

$$P_{\text{supply}}^{\text{SM}} = P_{\text{supply}}^{\text{CM}} - (N_{\text{RF}} - 1)\,P_0. \tag{11}$$

Thus, for the same transmission power, the total power consumption gain of SM, compared to other C-MIMO architectures, scales linearly with respect to the number of RF chains.

In the second scenario, we assume that every single antenna of a C-MIMO BS transmits the same amount of power as the single activated antenna of a BS employing SM. Similar to the previous analysis, using (9) and (10), we can show that

$$P_{\text{supply}}^{\text{SM}} = \frac{P_{\text{supply}}^{\text{CM}}}{N_{\text{RF}}}. \tag{12}$$

It is clear, from the previous analysis, that in both cases a multi-antenna BS employing SM requires less total power than a C-MIMO BS with the same number of transmit antennas.

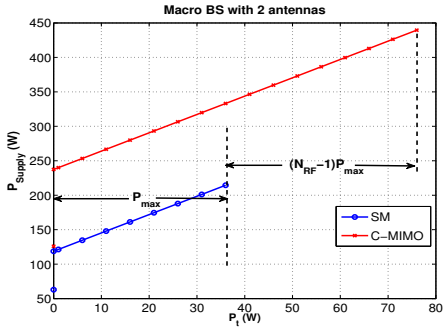


Fig. 3. P_{t} versus P_{supply} for a Macro BS with $N_t = 2$ transmit antennas (SOTA 2010).

In order to visualize the previous conclusion, we plot in Fig. 3 the total power consumption versus the transmitted power for a Macro BS with $N_t = 2$ antennas, using the power model in (8) and the parameters from Table I. Furthermore, Fig. 3 shows that a BS employing a C-MIMO architectures achieves higher transmission power, compared to a BS employing SM, resulting in the increase of the already high power consumption.

VI. BS CAPACITY WITH RESPECT TO THE TOTAL POWER CONSUMPTION

In the previous section, we show that for the same transmit power P_{t}, SM requires less power than a multiple RF chain scheme. Furthermore, we show that the power reduction at a BS using SM increases with respect to the number of RF chains employed in the C-MIMO architectures.

TABLE II
SIMULATION PARAMETERS.

Simulation Parameters	Values
BS Type	Macro
Power Model Parameters	SOTA 2010
Carrier Frequency	2 GHz
Path Loss Model	3GPP UMa, NLOS, [14]
Shadowing Standard Deviation	6 dB
Number of Channels	10000
Bandwidth	10 MHz
Operating Temperature	290 K
d (BS to user distance)	100 m

In this section, we compare different multi-antenna BS architectures in terms of capacity versus the total power consumption. Specifically, we compare the achievable ergodic capacity of SM [4], STBC [7], and MISO with only transmit diversity [12], [6].

The capacity for a STBC system with N_t transmit and N_r receive antennas, given a flat channel, is written as

$$C_{\text{STBC}} = WR \log_2\left(1 + \frac{P}{NN_t}\sum_{i=1}^{N_t}\|\mathbf{h}_i\|_2^2\right), \qquad (13)$$

where R is the rate of a certain STBC code. For complex transmit signals, the rate of STBC is $R = 1$ only for the special case of $N_t = 2$. Generally, $R = \frac{1}{2}$ for $N_t > 2$ with exception of two and and three transmit antennas, where $R = \frac{3}{4}$ [7], [13].

Furthermore, the capacity for a MISO system with only transmit diversity and N_t transmit antennas is formulated as [12], [6]

$$C_{\text{MISO}} = W \log_2\left(1 + \frac{P\|\mathbf{h}\|_2^2}{N}\right), \qquad (14)$$

where $\mathbf{h}$ is the MISO channel.

Using Monte Carlo simulations, we evaluate the ergodic capacity for a BS with $N_t = \{2, 8\}$ transmit antennas and a single user with one receive antenna. Given that the noise power is defined as $N = W\kappa\theta$, where κ is the Boltzmann constant and θ is the operating temperature in K, we simulate a scenario with parameters presented in Table II.

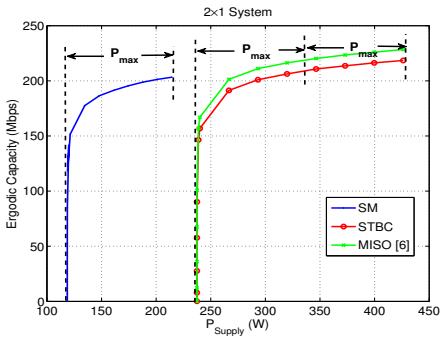


Fig. 4. Ergodic capacity versus power supply for a BS with $N_t = 2$.

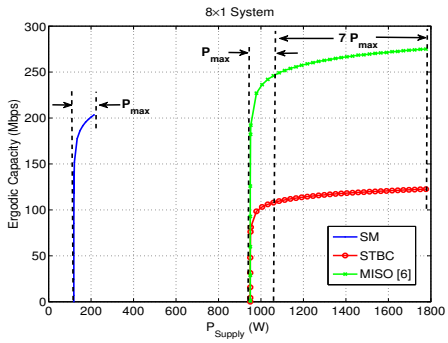


Fig. 5. Ergodic capacity versus power supply for a BS with $N_t = 8$.

The simulation results are presented in Fig. 4 and 5. Clearly, in all cases, a BS employing SM requires less total power supply P_{supply} in order to achieve the same capacity as C-MIMO architecture, when the transmitted power is less than P_{max}. More specifically, for a transmit antenna configuration $N_t = 2$, SM has a power reduction of 52.5 W compared to MISO and 79.5 W compared to STBC, at the level of SM's maximum achievable capacity (214.5 Mbps). This translates to 19% and 32% power reduction, respectively. It is clear from Fig. 4 that C-MIMO schemes achieve higher capacity than SM when their transmitted power is higher than P_{max}. However, an additional cost has to be paid in power consumption which reaches to 426 W for their maximum achievable capacity (i.e. 160% increased power consumption compared to SM).

Similar conclusions can be drawn from Fig. 5 except the fact that SM achieves higher capacity than STBC. In Fig. 5, SM has a capacity gain of 79.2 Mbps (39% increase) and a power consumption gain of 1573.5 W over STBC (90% decrease), at the maximum achievable capacity of STBC.

The combination of the capacity and power gain of SM, compared to STBC, stems from the fact that SM is a single RF chain scheme which achieves higher capacity for a number of transmit antennas greater than 2. As for the comparison of SM and MISO, the achievable capacity of MISO is higher than SM. The additional cost that has to be paid is 1573.5 W of increased power consumption (an increase of factor 9), compared to the power consumption of SM under the maximum achievable capacity.

VII. ANALYSIS OF ENERGY EFFICIENCY

In this section, we study the energy efficiency (in terms of Mbps/J) of the different MIMO schemes versus the achievable ergodic capacity. The definition of the energy efficiency we use in this work is the ratio of the transmitted bits and the total energy consumption. In order to quantify the fundamental limits of each MIMO architecture, we use the ergodic capacity. It can be proven that the energy efficiency is reduced to

$$EE(P_{\text{t}}^{\text{total}}) = \frac{C_{\text{ergodic}}(P_{\text{t}}^{\text{total}})}{P_{\text{supply}}(P_{\text{t}}^{\text{total}})}, \qquad (15)$$

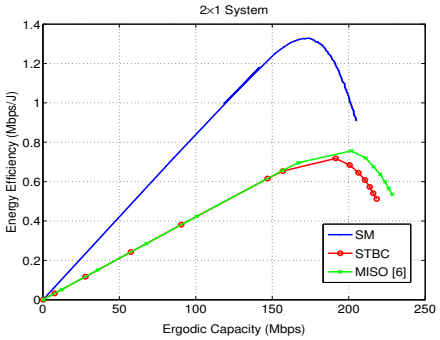


Fig. 6. Energy Efficiency versus Ergodic Capacity for a BS with $N_t = 2$.

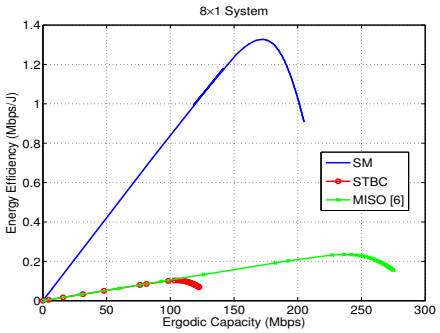


Fig. 7. Energy Efficiency versus Ergodic Capacity for a BS with $N_t = 8$.

where $P_{\text{t}}^{\text{total}}$ is the total RF power transmission.

Using Monte Carlo simulation results, in Fig 6 and 7, we plot the $EE(P_{\text{t}}^{\text{total}})$ of the different MIMO schemes versus the achievable ergodic capacity. In all cases, it is clear that SM is the most energy efficient scheme. This phenomenon can be explained by the fact that SM requires less power supply P_{supply} (as shown in Section VI) in order to transmit the same number of bits. In addition, it can be seen that there is an optimal operation point where the $EE(P_{\text{t}}^{\text{total}})$ is maximized for each scheme. This stems from the fact that after a given power transmission point, the capacity of all schemes does not change rapidly (Fig. 4 and 5).

VIII. Conclusions

In this paper, we study the total power consumption of a BS incorporating SM. Although SM is a MIMO scheme, its basic operating principle offers a single RF chain system implementation. Thus, an important reduction on power supply is achieved, compared to other multiple-RF chain MIMO architectures. Using information theoretic results, we show that a BS employing SM has an important power consumption gain compared to other MIMO scheme, retaining the same ergodic capacity. Especially for an antenna configuration $N_t > 2$, SM has a higher capacity than STBC, combined with a power consumption gain which scales in respect to the number of RF chains employed by STBC. SM can reduce overall power consumption by nine times for $N_t = 8$, when compared to C-MIMO. In addition, we illustrate the energy efficiency (bits/J) performance of all studied schemes and show that SM is the most energy efficient among them. Based on simulation results, we show that there is a transmission point where the energy efficiency is maximized.

Finally, the main concern of our future work will be the expansion of our study to different MIMO schemes and antenna configurations, combined with an analytical framework for the energy efficiency results.

Acknowledgement

This work has been funded by the Research Project GREENET (PITN-GA-2010-264759).

References

[1] E. Telatar, "Capacity of Multi-Antenna Gaussian Channels," *European Transaction on Telecommunication*, vol. 10, no. 6, pp. 585–595, November / December 1999.

[2] A. Molisch and M. Win, "MIMO Systems With Antenna Selection," *Microwave Magazine, IEEE*, vol. 5, no. 1, pp. 46 – 56, Mar 2004.

[3] L. Correia, D. Zeller, O. Blume, D. Ferling, Y. Jading, I. Go anddor, G. Auer, and L. Van Der Perre, "Challenges and Enabling Technologies for Energy Aware Mobile Radio Networks," *Communications Magazine, IEEE*, vol. 48, no. 11, pp. 66 –72, November 2010.

[4] M. Di Renzo, H. Haas, and P. M. Grant, "Spatial Modulation for Multiple-Antenna Wireless Systems: A Survey," *IEEE Commun. Mag.*, vol. 49, no. 11, pp. 182–191, Nov. 2011.

[5] G. Auer, V. Giannini, I. Godor, P. Skillermark, M. Olsson, M. Imran, D. Sabella, M. Gonzalez, C. Desset, and O. Blume, "Cellular Energy Efficiency Evaluation Framework," in *Vehicular Technology Conference (VTC Spring), 2011 IEEE 73rd*, May 2011, pp. 1 –6.

[6] G. J. Foschini and M. J. Gans, "On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas," *Wireless Personal Communications*, vol. 6, no. 6, pp. 311–335, 1998.

[7] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time Block Codes from Orthogonal Designs," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1456–1467, Jul. 1999.

[8] J. Jeganathan, A. Ghrayeb, and L. Szczecinski, "Spatial Modulation: Optimal Detection and Performance Analysis," *IEEE Commun. Lett.*, vol. 12, no. 8, pp. 545–547, 2008.

[9] Y. Yang and B. Jiao, "Information-Guided Channel-Hopping for High Data Rate Wireless Communication," *IEEE Commun. Lett.*, vol. 12, no. 4, pp. 225 –227, Apr. 2008.

[10] H. Holtkamp, G. Auer, and H. Haas, "On Minimizing Base Station Power Consumption," in *Proc. of the Vehicular Technology Conference (VTC)*, IEEE. San Francisco, CA, USA: IEEE, Sep. 5–8, 2011, pp. 1–5.

[11] P. Wolniansky, G. Foschini, G. Golden, and R. Valenzuela, "V-BLAST: an Architecture for Realizing very High Data Rates over the Rich-Scattering Wireless Channel," in *Unino Radio-Scientifique Internationale (URSI) Intern. Symp. on Signals, Systems, and Electronics (ISSSE)*, Sep. 29–Oct. 2, 1998, pp. 295–300.

[12] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[13] S. Sandhu and A. Paulraj, "Space-Time Block Codes: A Capacity Perspective," *Communications Letters, IEEE*, vol. 4, no. 12, pp. 384 –386, Dec. 2000.

[14] 3GPP, "Further Advancements for E-UTRA Physical Layer Aspects (Release 9)," 3GPP TR 36.814 V0.4.1 (2009-02), Sep. 2009. Retrieved Jun. 2, 2009 from www.3gpp.org/ftp/Specs/.

F.5 IEEE 46th Asilomar Conf. on Signals, Systems and Computers (ASILOMAR)

A Power Saving Dual-Hop Architecture Based on Hybrid Spatial Modulation

Athanasios Stavridis*, Sinan Sinanović*, Marco Di Renzo†, and Harald Haas*

*Institute for Digital Communications
Joint Research Institute for Signal and Image Processing
School of Engineering
The University of Edinburgh
{a.stavridis, s.sinanovic, h.haas}@ ed.ac.uk

†Laboratory of Signals and Systems (L2S),
French National Center for Scientific Research (CNRS)
École Supérieure d'Électricité (SUPÉLEC),
University of Paris–Sud XI (UPS)
marco.direnzo@ lss.supelec.fr

Abstract **In this paper, we propose a novel Dual-Hop Decode-and-Forward (DF) architecture based on Multiple-Input Multiple-Output (MIMO) Zero Forcing (ZF) precoding and Spatial Modulation (SM) that employs a centralized or a distributed detection algorithm at the Relay Nodes (RNs). Using Tikhonov Regularization (TR) we form a precoding techni ue that turns out to reduce significantly the transmitted power at the Source Node (SN) under certain scenarios. Moreover, the use of TR enables our scheme to communicate in channels with low and medium correlation at the first hop without a Bit Error Rate (BER) penalty. Finally, we extend our scheme in order to take into account realistic Channel State Information (CSI) at the the Source Node (SN).**

I. Introduction

SM is a relatively new single Radio Frequency (RF) chain Multiple-Input Multiple-Output (MIMO) scheme [1]. The basic concept of SM is to split the transmitted bits into two sequences. The first sequence is modulated by a conventional constellation point like M-ary Quadrature Amplitude Modulation (QAM), whereas the second sequence is used for the selection of only one transmitting antenna. Thus, SM entirely avoids Inter-Antenna Synchronization (IAS) and Inter-Channel Interference (ICI). At the receiver, SM jointly detects the transmitted constellation point and the transmitting antenna index [2]. A notable attribute of SM is its high Energy Efficiency (EE). In [3], we demonstrate that the single RF chain configuration of SM offers a significant EE gain compared to conventional MIMO schemes like [4, 5].

Another recently proposed SM-like scheme is Receive-Spatial Modulation (R-SM) [6, 7]. R-SM is the reciprocal of SM. Instead of activating only one transmit antenna as SM, R-SM employs ZF precoding and the appropriate signal formation in order to aim at only one receiving antenna. A low complexity non-coherent Maximum Likelihood (ML) receiver jointly detects the transmitted constellation point and the index of the receiving antenna [7].

Motivated by the EE that relays networks might offer [8] and the potential of SM/R-SM, we propose a non-cooperative relay architecture based on SM/R-SM. Our architecture includes a multi-antenna BS as a SN, multiple RNs which are single antenna elements (with or without backhaul connection), and a Destination Node (DN) which is a single or a multi-antenna element.

Using a half-duplex DF protocol we propose a scheme that achieves information conveyance using R-SM [7] in the first slot and SM [1] in the second slot. In this way, a Dual Hop-Hybrid Spatial Modulation (DH-HSM) system is formed. The term Hybrid Spatial Modulation stems from the fact that our architecture employs R-SM in the first hop and SM in the second hop. DH-HSM conveys information by extending the novel transmission mechanism of SM to a distributed framework.

In the first phase, the SN aims at only one RN using ZF precoding under Channel State Information at the Transmitter (CSIT). In the second phase, only the receiving RN from the previous phase is activated forming a SM modulation scheme from the RN to the DN. Clearly, the index of the receiving/activating RN can be considered as an additional mechanism of information transmission except from the conventional constellation point.

As a consequence, DH-HSM offers an energy consumption gain at the RNs due to the activation of only one RN during the signaling period of the second slot. The value of this property is even more significant when the RNs are remote terminals due to the battery life extension.

It is well known that ZF precoding introduce three difficulties for a real system implementation. The first difficulty is originated from the fact that precoding requires CSIT. Thus, initially we study the proposed architecture under Perfect CSIT (P-CSIT) and then we extend our scheme in order to take into account realistic scenarios of CSIT. The other difficulties that ZF precoding may face are: (a) channel correlation and (b) the high transmitted power. Hence, in this paper we employ TR from optimization theory that aims to resolve these issues.

The rest of the paper is organized as follows: Section II presents the system model. Section III proposes a distributed detection algorithms that can be employed at the relays. A precoding scheme that offers a transmit power gain at the SN is presented in Section IV. Section V studies the proposed architecture under realistic scenarios of CSIT at the SN. The simulation results of DH-HSM are presented in Section VI. Finally, we conclude this paper in Section VII.

Notation: In the following, lowercase bold letters denote vectors and uppercase bold letters denote matrices. $(\cdot)^T$, $(\cdot)^H$ denotes transpose and Hermitian transpose, respectively. $\mathrm{Re}\{\cdot\}$

stands for the real part of a complex scalar or matrix. $\mathrm{E}[\cdot]$ is the mean value of a RV. Finally, $\|\cdot\|_2$ and $\|\cdot\|_F$ stands for the Euclidean and Frobenius norm, respectively.

II. System Model

Fig. 1 depicts the system model of the proposed DH-HSM relay architecture. We consider a BS equipped with N_t antennas which acts as the SN. Furthermore, we assume that our architecture includes N_R single antenna RNs connected with an error-free backhaul link. Later in Section III, we extend our architecture in a distributed framework where there is no backhaul connection between the RNs. The DN can be considered as a single or multi-antenna N_D node.

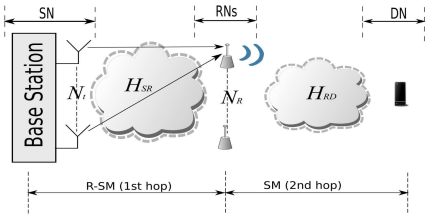


Fig. 1. System Model of DH-HSM.

In addition, we assume that there is no direct link between the SN and the DN. The SN can acquire CSI using either the reciprocity principle when is applicable or using a low-rate feedback link from the relays. The DN can acquire CSI using a training sequence transmitted from the relays. In this Section, we assume P-CSIT at DN for the sake of system presentation and in Section V we modify our architecture in order to operate under different practical CSI scenarios at the SN.

If we configure our system with $N_t \geq N_R$ and N_R is a power of 2, the received signal at the RNs during the first hop can be written in a matrix form as,

$$\mathbf{y}_R = \mathbf{H}_{SR}\mathbf{P}\mathbf{x}(i, s_k) + \mathbf{w}_{SR}. \tag{1}$$

In (1), $\mathbf{y}_R$ is a $N_R \times 1$ vector whose i-th element is the received signal at the i-th RN, $\mathbf{H}_{SR}$ is a $N_R \times N_t$ matrix representing the wireless channel between every transmitting antenna and every RN, $\mathbf{P}$ is the $N_t \times N_R$ precoding matrix, $\mathbf{x}(i, s_k)$ is the transmitted signal vector, and $\mathbf{w}_{SR} \in \mathcal{C}^{N_r}$ is the zero mean vector of complex additive Gaussian noise with independent identically distributed (i.i.d.) elements of variance of $\sigma^2_{w_{SR}}$. The transmitted vector can be written as $\mathbf{x}(i, s_k) = \mathbf{e}_i s_k$, where the vector $\mathbf{e}_i = [0, ..., 0, 1, 0, ...0]^T$, $i = 1, ..., N_R$, has length of N_R and all of its elements are zero except the one at the i-th position which has the value of 1. $s_k \in \{s_1, ..., s_M\}$ represents the transmitted symbol selected from a conventional constellation.

The design of the MIMO precoding matrix $\mathbf{P}$ is extensively studied in the literature. Examples of designing criterion for the precoding matrix $\mathbf{P}$ include the minimization of the Pairwise Error Probability (PEP) or the maximization of the receive Signal-to-Noise Ratio (SNR) [9]. In this paper, we employ the ZF precoding method, where $\mathbf{H}_{SR}\mathbf{P} = \mathbf{I}_{N_R, N_R}$. We adopt this method because we are interested on the total elimination of ICI. Using the Singular Value Decomposition (SVD) the precoding matrix is formed as $\mathbf{P} = \mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^H$, where $\mathbf{H}_{SR} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$. It is generally known that ZF precoding may increase the transmit power in order to overcome deep fades creating a number of difficulties in real system implementation. In order to overcome this drawback, in Section IV we propose a precoding method based on TR which among others trades off between the residual ICI and the transmitted power.

Under the assumption of ZF precoding using P-CSIT at the SN, the observation at each of the RNs is given as,

$$\begin{aligned} y_{Rj} &= s_k + w_{SRj}, \quad j = i, \\ y_{Rj} &= w_{SRj}, \quad\quad\ j \neq i, \end{aligned} \tag{2}$$

$j = 1, ..., N_R$. As it can be seen in (2), only one RN receives the transmitted symbol s_k degraded from Gaussian noise and all other RNs are facing only Gaussian noise. This transmission scheme allows to map $k_1 = \log_2 N_R$ bits on the selection of the index of the i-th receiving RN (i corresponds also to the position of the non-zero element in $\mathbf{e}_i$) and $k_2 = \log_2 M$ bits on the transmitted constellation symbol s_k.

Due to the fact that the RNs communicate via a backhaul-link, we employ the centralized low-complexity non-coherent ML detector proposed in [7]

$$(\hat{i}, \hat{s}_k) = \arg\min_{i, s_k} |s_k|^2 - 2\mathrm{Re}\{y_{Ri}^* s_k\}. \tag{3}$$

During the second slot, we apply a SM-like modulation scheme at which only the receiving RN from the previous slot is activated acting as a single antenna in a conventional SM MIMO system. In this way, the system equation for the RNs-DN link is given as,

$$\mathbf{y}_D = \mathbf{H}_{RD}\mathbf{x}(\hat{i}, \hat{s}_k) + \mathbf{w}_{RD}, \tag{4}$$

where $\mathbf{H}_{RD}$ stands for the $N_D \times N_R$ wireless channel from the RNs to DN, $\mathbf{x}(\hat{i}, \hat{s}_k) = \mathbf{e}_{\hat{i}}\hat{s}_k$ is the $N_R \times 1$ vectorized detected symbol from the previous slot. All the elements of $\mathbf{x}(\hat{i}, \hat{s}_k)$ are zero except the one at the $\hat{i}$-th position which is set equal to $\hat{s}_k$. This modulation scheme corresponds to the scenario where only the receiving RN is relaying the transmitted conventional constellation point. Moreover, $\mathbf{w}_{RD} \in \mathcal{C}^{N_D}$ is the i.i.d. additive complex Gaussian noise with $\mathbf{w}_{RDi} \sim \mathcal{CN}(0, \sigma^2_{w_{RD}})$. Finally, $\mathbf{y}_D$ stands for the received signal at the DN. The optimal ML detection for SM is given in [2]

$$(\check{i}, \check{s}_k) = \arg\min_{i, s_k} \|\mathbf{y}_D - \mathbf{H}_{RD}\mathbf{e}_i s_k\|_2^2. \tag{5}$$

Hence, we form a Dual-Hop architecture in which a total of $k = \log_2 N_R + \log_2 M$ bits are transmitted using two symbol periods.

III. Extension to a Distributed Relays Architecture

In some scenarios there is no backhaul-link between the RNs. This is the fact when the delay requirement does not

allow communication between the RNs or when the RNs are remote terminals (like cell phones). This Section extends the proposed centralized relay architecture from Section II to a distributed one with the employment of a distributed detection algorithm at the RNs. All the other configurations of the proposed architecture remain the same.

During the first hop the received signal at at the j-th RN is described in (2). It is straightforward to see that if we set $s_0 = 0$, (2) forms a $(M+1)$-ary Hypothesis Test (HT)

$$H_i: \quad y_{Rj} = s_i + w_{SRj}, \quad i = 0, ..., M, \tag{6}$$

where $s_i \in \{s_0, s_1, ..., s_M\}$. Clearly, the H_0 hypothesis ($s_0 = 0$) corresponds to the case where the j-th RN is not the receiving node in the first hop, thus it is not the activated node in the second hop. On the other hand, all the other $H_1, ..., H_M$ hypotheses correspond to the case where the j-th node is the receiving/activating RN and the relayed QAM symbol is $s_1, ..., s_M$, respectively.

In this case, the distributed ML detector at each RN is reduced to the usual minimum distance rule

$$(\hat{s}_i) = \arg \min_{s_i \in \{s_0, s_1, ..., s_M\}} |y_{Rj} - s_k|, \tag{7}$$

which is a non-coherent detector. As we can see, the computational complexity of (7) at each RN is just $M+1$ complex subtractions followed by the operation of finding the minimum value of the detector's evaluation.

IV. R-SM using Tikhonov Regularization at the Source Node

ZF precoding has the major disadvantage of increasing the transmitted power when the wireless channel is in deep fade. In order to overcome this major disadvantage of ZF precoding, we employ the well known TR method from optimization theory, which results into closed form solution [10]. TR formulates the computation of every column $\mathbf{p}_i$ of the precoding matrix $\mathbf{P}$ as the solution of the following quadratic minimization problem

$$\min_{\mathbf{p}_i} \|\mathbf{H}_{SR}\mathbf{p}_i - \mathbf{e}_i\|_2^2 + \delta\|\mathbf{p}_i\|_2^2, \tag{8}$$

where $\mathbf{e}_i$ is the i-th column of the ideal ZF impulse response and $\delta > 0$ is the regularization parameter [10]. The analytical solution has the following form

$$\mathbf{p}_i = \left[\mathbf{H}_{SR}^H\mathbf{H}_{SR} + \delta\mathbf{I}\right]^{-1}\mathbf{H}_{SR}^H\mathbf{e}_i \tag{9}$$

and can be reached using the gradient condition.

TR is the most famous regularization method that penalizes big values of $\|\mathbf{p}_i\|_2^2$ [10]. Hence, by using TR we impose an indirect regularized transmit power constraint on each $\mathbf{p}_i$ ($\|\mathbf{P}\|_F^2 = \sum_{i=1}^{N_R}\|\mathbf{p}_i\|_2^2 \leq p_{TR}$). Generally, values of δ close to 0 produce precoding vector with the same transmit power as the ZF precoder. On the other hand when δ is increased the total transmit power is decreased with the concurrent increase of the residual ICI at the receive antennas. Finally, due to the fact that $\mathbf{H}_{SR}^H\mathbf{H}_{SR} + \delta\mathbf{I} \succ \mathbf{0}$ for any $\delta > 0$, TR precoding does not pose any rank restriction on the channel matrix $\mathbf{H}_{SR}$ [10]. This valuable characteristic of TR gives the opportunity to the proposed precoder to face correlated channel.

V. Precoding Using Practical CSIT at the Source Node

In realistic systems providing the transmitter with P-CSIT is almost impossible. Thus, in literature, there is a number of precoding methods which employ limited feedback in the form of channel quantization or in the form of statistical CSIT [11], [9]. In this Section, we modify the precoding scheme between the SN and the RNs in order to take into account scenarios of limited feedback in the form of Quantized CSIT (Q-CSIT), Statistical CSIT (S-CSIT), and Worst-Case CSIT (WC-CSIT).

Motivated from [12], even though quantization is out of the scope of our work, we employ a simple Scalar Quantizer (SQ) that can be used in the limited feedback from the RNs to the SN. In this cases, the real and imaginary part of every element $h_{i,j}$ of $\mathbf{H}_{SR}$ is quantized separately using a SQ of K bits with a dynamic range of $[d_{\min}, d_{\max}]$, where $d_{\min}$ and $d_{\max}$ is the minimum and the maximum element, respectively, of the real or imaginary part of the channel matrix. A comprehensive description of sophisticated methods on channel quantization or quantized MIMO precoding can be found at [11].

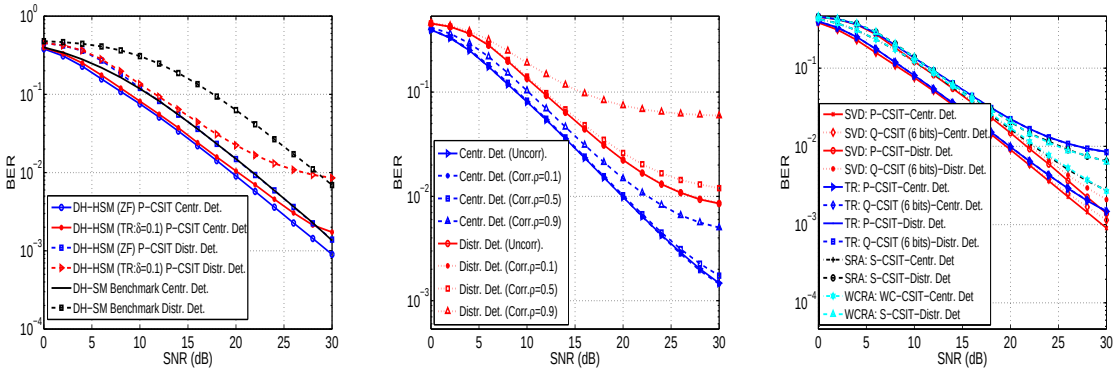
An alternative realistic scenario of CSI at SN is a slow time varying statistical feedback of the channel matrix like the mean value or the autocorrelation matrix of the channel [9]. In this case the channel matrix $\mathbf{H}_{SN}$ is written as $\mathbf{H}_{SN} = \overline{\mathbf{H}}_{SN} + \widetilde{\mathbf{H}}_{SN}$, where $\overline{\mathbf{H}}_{SN}$ represents the long term channel evolution which can be accurately acquired by the transmitter, whereas $\widetilde{\mathbf{H}}_{SN}$ denotes the channel uncertainty for which only statistical or worst-case knowledge can be assumed accurate. Usually $\widetilde{\mathbf{H}}_{SN}$ is modeled as Zero Mean Circular Symmetric Complex Gaussian (ZMCSCG) matrix with i.i.d. elements of variance $\sigma^2_{\widetilde{\mathbf{H}}_{SN}}$.

In this paper, we employ the precoding methods initially proposed in [13] for MIMO frequency selective channels and extended in [7] for the R-SM transmission scheme. The first method of [7] is based on Statistical Robust Approximation (SRA) and every column of the precoding matrix is written as $\mathbf{p}_i = \left[\overline{\mathbf{H}}_{SN}^H\overline{\mathbf{H}}_{SN} + \mathrm{E}\left[\widetilde{\mathbf{H}}_{SN}^H\widetilde{\mathbf{H}}_{SN}\right]\right]^{-1}\overline{\mathbf{H}}_{SN}^H\mathbf{e}_i$. The second method is based on Worst-Case Robust Approximation (WCRA) and every column of the precoding matrix is written as $\mathbf{p}_i = \left[\overline{\mathbf{H}}_{SN}^H\overline{\mathbf{H}}_{SN} + \beta\mathbf{I}\right]^{-1}\overline{\mathbf{H}}_{SN}^H\mathbf{e}_i$, where β is a constant which depends on the value of $\|\widetilde{\mathbf{H}}_{SN}\|_2$ [7].

It should be noted that both SRA and WCRA are forms of TR with a very good behavior under correlated channel conditions combined with a transmit power gain [7].

VI. Simulation Results

In this Section, we present Monte Carlo simulation results that demonstrate the performance of the proposed DH-HSM architecture under different CSIT scenarios at the SN, both for the centralized (Section II) and distributed (Section III) detection algorithms. The system configuration is $N_t = 4$ transmit antennas at the BS, $N_R = 4$ single-antenna RNs and a single-antenna DN. Furthermore, the transmitted constellation



(a) DH-HSM versus benchmark (DH-SM).

(b) DH-HSM TR ($\delta = 0.1$) with P-CSIT under different correlation scenarios (different values of ρ).

(c) DH-HSM with imperfect CSI.

Fig. 2. BER performance of DH-HSM (centralized and distributed detection) for the symmetric case.

points s_k are selected from a 4-QAM constellation with an average transmit power equal to 1. Hence, a total of $k = 4$ bits are conveyed from the SN to the DN into two symbol periods (two hops).

We study the symmetrical case where both the SN-RNs channel $\mathbf{H}_{SR}$ and RNs-DN channel $\mathbf{H}_{RD}$ are uncorrelated ZMCSCG matrices with zero mean i.i.d. elements of variance equal to 1. Furthermore, in order to evaluate our system performance under the existence of correlations at the SN-RNs channel we employ the well known Kronecker correlation model with spatial correlation matrices generated from an exponential model. Due to the space limitation we refer the reader to [7] where there is a comprehensive description of the correlation model we use. Finally, the receive SNR for the first hop is defined as $\text{SNR}_{\text{dB}}^{SR} = 10\log\frac{P_s\|\mathbf{P}\mathbf{H}_{SR}\|_F^2/N_R}{\sigma_{w_{SR}}^2}$ and the receive SNR for the second hop is defined as $\text{SNR}_{\text{dB}}^{RD} = 10\log\frac{P_s\|\mathbf{H}_{SR}\|_F^2/N_R}{\sigma_{w_{RD}}^2}$. For both cases we set $P_s = 1$ and $\text{SNR}^{SR} = \text{SNR}^{RD}$ (symmetrical case).

Fig. 2(a) compares the BER performance of DH-HSM under P-CSIT at the SN and under centralized or distributed detection at the RNs. In addition, for the sake of comparison, we include as a benchmark the BER performance of Dual Hop-Spatial Modulation (DH-SM) when a centralized [14] or a distributed ML detector is employed. The employed distributed detector is similar to (7) and minimizes the distance rule $|y_{R_j} - h_{SR_{i,j}}s_k|$. $h_{SR_{i,j}}$ denotes the (i,j) element of $\mathbf{H}_{SR}$. In both hops, SNR is defined as the second hop of DH-HSM because this is the receive SNR of SM.

Inspecting Fig. 2(a), we see that when DH-HSM employs distributed detection at the RNs there is a performance penalty compared to the case of centralized detection. Under ZF precoding, this penalty is between 2 and 3 dB. For the case of TR precoding, the performance loss is 3 dB for SNRs ≤ 14 dB. Furthermore, as the value of SNR increases, the performance loss also increases due to the fact that in high SNRs the main degradation factor for TR precoding is the residual ICI which affects the distributed detector in a negative way. As regard the comparison of DH-HSM under ZF and TR precoding, we see that both schemes have almost the same performance for SNRs ≤ 16 dB. In high SNRs, ZF precoding overcomes TR precoding again due to the residual ICI.

Concerning the comparison between DH-HSM and DH-SM, DH-HSM has better BER performance than DH-SM due to the employment of the precoding scheme at the first hop. There is an exception for SNRs > 28 dB where DH-SM overcomes DH-HSM with TR precoding. Finally, it is very interesting to note that DH-HSM combined with ZF precoding and distributed detection has the same performance as DH-SM with centralized detection for SNRs higher than 8 dB.

Fig. 2(b) demonstrates the BER performance of DH-HSM under different correlation scenarios at the first hop, when TR precoding is employed with a regularization parameter $\delta = 0.1$. We assume that the SN-RNs channel $\mathbf{H}_{SR}$ is correlated according the Kronecker correlation model employed in [7]. In this correlation model, ρ stands for the correlation factor used in the exponential correlation model between different antennas [7]. In addition, we assume that the RNs-DN channel $\mathbf{H}_{RD}$ is uncorrelated which corresponds to the scenario that the RNs are located far apart. Inspecting Fig. 2(b) we see that for low ($\rho = 0.1$) and median ($\rho = 0.5$) correlation there is no performance difference from the uncorrelated case. For high correlation ($\rho = 0.9$) and for the case of centralized detection the performance difference ranges from a fraction of a dB in low SNRs up to 7 dB in high SNRs. The same phenomenon can be observed under distributed detection at the RNs with a more diminishing performance in high SNRs.

Moreover, Fig. 2(c) presents the performance of DH-HSM under practical scenarios of Q-CSIT (using the simple quantizer described in Section V), S-CSIT, and WC-CSIT at the SN. It is possible to have no performance difference between Q-CSIT and P-CSIT if the number of the feedback bits is big

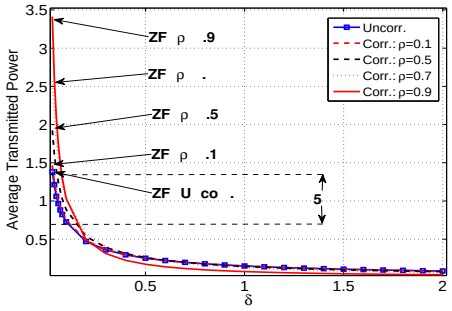


Fig. 3. Average transmit power at the SN versus δ, when TR precoding is employed under P-CSIT, for different correlation scenarios.

enough (for our simulations, we have used 6 bits separately for the real and the imaginary part of every channel coefficient). Furthermore, in high SNRs we see that for 6 bits, TR has better behavior compared to ZF under Q-CSIT.

In agreement with [7], for the case of S/WC-CSIT we assume that the elements of $\overline{\mathbf{H}}$ are refreshed every 100 symbols and follow an i.i.d. zero mean complex Gaussian distribution with variance 0.99. The elements of $\widetilde{\mathbf{H}}$ are refreshed every symbol period and follow the same distribution with variance 0.01. As it can be seen from Fig. 2(c), when SRA or WCRA precoding is employed at the SN and centralized detection at the RNs, there is a penalty of 4 dB compared to the scenario of ZF precoding under P-CSIT. In addition, when distributed detection is employed at the RN, a small performance degradation is observed compared to the scenario of centralized detection at RNs.

Fig. 3 presents the average transmitted power ($\mathrm{E}[\|\mathbf{P}\mathbf{e}_i s_k\|_2^2]$) of the first hop versus δ (regularization parameter), when TR precoding is employed with P-CSIT and under different correlation scenarios. Although Fig. 3 demonstrates the average transmitted power during the first hop, this figure is insightful of the whole system average transmitted power because the average transmitted power during the second hop is stable and depends on the normalized constellation power. As the value of δ is increased, the average transmitted power is reduced under all correlation scenarios. In contrast, when $\delta \to 0$, TR precoding reduces to ZF precoding. As a consequence, TR precoding offers a significant transmit power gain compared to ZF precoding, especially when the wireless channel is correlated. When δ is increased, the residual ICI at each RN is also increased resulting in a diminishing BER performance (due to space limitation we do not include the BER performance with respect to δ). Thus, there is a trade-off between the BER performance and the reduced average transmitted power. Though, as it can be seen from Fig. 2(a) and Fig. 3, for a value of $\delta = 0.1$ TR precoding has almost the same BER performance as ZF precoding for SNRs lower than 20 dB, utilizing at the same time only the half average transmitted power. Finally, it is interesting to note that for values of δ greater than 0.4, in practice the average transmitted power is stabilized. Hence, there is no point of using values of δ greater than 0.4.

VII. Conclusions

In this paper, we proposed a novel dual hop relaying architecture based on ZF precoding and SM called DH-HSM. We have studied its performance under idealistic and realistic CSI at the SN and we have shown that under sufficient feedback DH-HSM operates closely to the ideal scenario of P-CSIT at the SN. A transmit power gain has been achieved at the SN with the employment of TR precoding. TR precoding offers the ability to our scheme to cope with SN-RNs correlated links. In addition, we proposed a distributed detection algorithm which eliminates the communication between the RNs with a small BER penalty. Finally, under the correct detection at the RNs, the activation of only one RN during the second slot results in the reduction of the energy consumption at the RNs.

Acknowledgement

This work has been funded by the Research Project GREENET (PITN-GA-2010-264759).

References

[1] M. Di Renzo, H. Haas, and P. M. Grant, "Spatial Modulation for Multiple-Antenna Wireless Systems: A Survey," *IEEE Commun. Mag.*, vol. 49, no. 11, pp. 182–191, Nov. 2011.

[2] J. Jeganathan, A. Ghrayeb, and L. Szczecinski, "Spatial Modulation: Optimal Detection and Performance Analysis," *IEEE Commun. Lett.*, vol. 12, no. 8, pp. 545–547, 2008.

[3] A. Stavridis, S. Sinanovic, M. Di Renzo, H. Haas, and P. Grant, "An Energy Saving Base Station Employing Spatial Modulation," in *2012 IEEE 17th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, September 2012.

[4] G. J. Foschini and M. J. Gans, "On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas," *Wireless Personal Communications*, vol. 6, no. 6, pp. 311–335, 1998.

[5] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time Block Codes from Orthogonal Designs," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1456–1467, Jul. 1999.

[6] L.-L. Yang, "Transmitter Preprocessing Aided Spatial Modulation for Multiple-Input Multiple-Output Systems," in *Vehicular Technology Conference (VTC Spring), 2011 IEEE 73rd*, May 2011, pp. 1 –5.

[7] A. Stavridis, S. Sinanović, M. Di Renzo, and H. Haas, "Transmit Precoding for Receive Spatial Modulation Using Imperfect Channel Knowledge," in *GreeNet Workshop of the Vehicular Technology Conference (VTC Spring)*, IEEE. Yokohama, Japan: IEEE, May 6–9, 2012.

[8] D. Feng, C. Jiang, G. Lim, L. Cimini, Jr., G. Feng, and G. Li, "A Survey of Energy-Efficient Wireless Communications," *Communications Surveys Tutorials, IEEE*, vol. PP, no. 99, pp. 1 –12, 2012.

[9] M. Vu and A. Paulraj, "MIMO Wireless Linear Precoding," *Signal Processing Magazine, IEEE*, vol. 24, no. 5, pp. 86 –105, Sept. 2007.

[10] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[11] D. Love, R. Heath, V. Lau, D. Gesbert, B. Rao, and M. Andrews, "An Overview of Limited Feedback in Wireless Communication Systems," *Selected Areas in Communications, IEEE Journal on*, vol. 26, no. 8, pp. 1341–1365, October 2008.

[12] M. Xu, D. Guo, and M. Honig, "MIMO Precoding with Limited Rate Feedback: Simple Quantizers Work Well," in *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE*, Dec. 2009.

[13] A. Stavridis, S. Karachontzitis, and K. Berberidis, "Bezout-Based Robust Precoding for MIMO Frequency Selective Channel Using Imperfect Channel Knowledge," in *Proc. of the 19th European Signal Processing Conference (EUSIPCO)*, Barcelona, Spain, Aug. 2011.

[14] N. Serafimovski, S. Sinanovic, M. Di Renzo, and H. Haas, "Dual-hop Spatial Modulation (Dh-SM)," in *Proc. of the Vehicular Technology Conference (VTC Spring)*, IEEE. Budapest, Hungary: IEEE, May 15–18, 2011, pp. 1–5.

F.6 IEEE 78th Vehicular Technology Conference (VTC 2013-Fall)

Energy Evaluation of Spatial Modulation at a Multi-Antenna Base Station

Athanasios Stavridis*, Sinan Sinanović‡, Marco Di Renzo§, and Harald Haas*

*Institute for Digital Communications
Joint Research Institute for Signal and Image Processing
School of Engineering
The University of Edinburgh
Edinburgh, UK
{a.stavridis, h.haas}@ed.ac.uk

‡School of Engineering and Built Environment
Glasgow Caledonian University
Glasgow, UK
Sinan.Sinanovic@gcu.ac.uk

§Laboratory of Signals and Systems (L2S)
French National Center for Scientific Research (CNRS)
École Supérieure d'Électricité (SUPÉLEC)
University of Paris–Sud XI (UPS)
Paris, France
marco.direnzo@lss.supelec.fr

***Abstract*—In this paper, we aim to study the Energy Efficiency (EE) of Spatial Modulation (SM) at different Base Stations (BSs) taking into account the total power consumption. Compared to conventional Multiple-Input Multiple-Output (MIMO) schemes, SM benefits from a single Radio Frequency (RF) chain which results in decreased power supply (W), higher EE (Mbits/J), and reduced complexity. Using the fundamental limits of Shannon capacity, we show that SM achieves a range of average data rates with only a fraction, which can be as low as 24% for four transmit antennas, of the total power supply of conventional MIMO. In addition, we demonstrate that the EE of the studied schemes is maximized for a certain average data rate and that SM achieves the highest EE among them. Finally, we note that a BS employing SM can be up to 67% more energy efficient compared to a BS under a conventional MIMO transmission scheme, for four transmit antennas.**

I. Introduction

During the last fifteen years, research on Multiple-Input Multiple-Output (MIMO) wireless communication has demonstrated that multi-antenna communication is a promising solution for spectrally efficient future wireless networks. The use of multi-antenna elements, both at the transmitter and receiver, is shown to offer increased system capacity without the need for extra radio resources [1, 2].

MIMO communication includes a variety of transmission schemes. For example, the well known Vertical-Bell Laboratories Layered Space-Time (V-BLAST) algorithm is designed to offer high data rates using spatial multiplexing [3], while Space-Time Block-Coding (STBC) is designed to offer reliable data transfer [4]. Unfortunately, multi-antenna elements face a major disadvantage of requiring multiple Radio Frequency (RF) chains. RF chains are expensive circuits that do not follow Moore's law [5]. Thus, the real system implementation becomes expensive and often impractical. In addition, RF chains are electronic circuits that may reduce the Energy Efficiency (EE) of the wireless system. RF chains include Power Amplifiers (PAs) which are responsible for 50-80% of the total power consumption in the transmitter [6]. Hence, there is a significant interest in new MIMO schemes that do not require multiple RF chains. An example of such a scheme is Spatial Modulation (SM) [7].

Unlike Conventional MIMO (CMIMO) schemes, SM exploits multiple transmit antennas in a way that only a single RF chain is required. During the signaling period, the bits to be transmitted are divided into two blocks. The first block is encoded using a conventional signal constellation diagram such as Quadrature Amplitude Modulation (QAM), whereas the other block is used for the selection of the single transmitting antenna, which is part of the orthogonal spatial constellation diagram. Well known advantages of SM include the avoidance of Inter-Antenna Synchronization (IAS) at the transmitter and Inter-Channel Interference (ICI) at the receiver [7]. In addition, the detection at the receiver is performed using a low-complexity Maximum Likelihood (ML) detector [8].

In this paper, using the Energy Aware Radio and neTwork tecHnology (EARTH) power model (which relates the RF transmit power to the total power consumption of a Base Station (BS) [9]), we show that for the same RF transmit power, the total power saving of SM scales linearly with the number of RF chains required by CMIMO. The main contribution of this paper is the average data rate and EE evaluation of SM under different type of BSs (macro, micro, pico and femtocell BS) and the comparison with STBC, Multiple-Input Single-Output (MISO) with only transmit diversity, and MIMO. Using the fundamental limits of Shannon capacity, we show that for all types of BS, SM achieves a range of average data rates with significantly less power consumption compared to the following benchmark systems: MISO with only transmit diversity [1], STBC [4], and MIMO [2]. In addition, we show that the benchmark systems are up to 67% less energy efficient compared to SM. Finally, we demonstrate that all MIMO transmission schemes (including SM) achieve an average data rate where the EE, in terms of Mbits/J, is maximized.

The remainder of this paper is organized as follows: Section II presents SM and its information theoretic analysis. The power model of a BS under different MIMO transmission schemes is introduced in Section III. Section IV provides the average data rate and EE results for SM and the benchmark systems. Finally, we conclude this paper in Section V.

978-1-4673-6187-3/13/$31.00 ©2013 IEEE

Notation: In the following, lowercase bold letters denote vectors and uppercase bold letters denote matrices. Furthermore, a complex Gaussian distribution with mean m and variance σ_C^2 is represented as $\mathcal{CN}(m, \sigma_C^2)$, where its real and imaginary parts are independent and identically distributed (i.i.d.) Gaussian Random Variables (RV) with distribution $N(m, \frac{\sigma_C^2}{2})$. Finally, $\|\cdot\|_2$ and $\|\cdot\|_F$ stand for the Euclidean and Frobenius norm, respectively.

II. Spatial Modulation and Its Information-Theoretic Analysis

A. System Model

We consider a point-to-point MIMO system with N_t transmit antennas and N_r receive antennas. We further assume that only the receiver has perfect channel knowledge. When the wireless channel is flat, the general MIMO system equation is written as,

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}, \tag{1}$$

where $\mathbf{y}$ denotes the received signal vector of length N_r, $\mathbf{H}$ is a $N_r \times N_t$ matrix denoting the MIMO channel, $\mathbf{x}$ is the $N_t \times 1$ transmitted vector, and $\mathbf{w} \in \mathcal{C}^{N_r}$ is the i.i.d. additive complex Gaussian noise with $w_i \sim \mathcal{CN}(0, N)$. The basic concept of SM is the transmission of a standard symbol form the single activated transmit antenna. Thus, if we configure the number of transmit antennas to be a power of two, the transmitted signal vector is formulated as $\mathbf{x} = \mathbf{e}_i s_k$, where $\mathbf{e}_i = [0, ..., 0, 1, 0, ..., 0]^T$, $i = 1, ..., N_t$, is a $N_t \times 1$ vector whose elements are zero except the i-th element which has the value of 1. Clearly, the position of the non-zero element of $\mathbf{e}_i$ corresponds to the position of the activated antenna. In addition, s_k is a standard symbol like M-ary QAM or Phase Shift Keying (PSK). In this case, a total of $k = k_1 + k_2$ bits are transmitted per symbol period. The first $k_1 = \log_2 N_t$ bits are used for the selection of the single activated antenna, whereas the other $k_2 = \log_2 M$ bits are used for the selection of the transmitted symbol s_k.

The optimal ML detector of SM is proposed in [8]. This detector is formulated as,

$$(\hat{i}, \hat{s}_k) = \arg\min_{i, s_k} \|\mathbf{y} - \mathbf{H}\mathbf{e}_i s_k\|_2^2. \tag{2}$$

As it can be seen, the index of the activated antenna and the transmitted standard symbol s_k are jointly detected.

B. Capacity of Spatial Modulation

The capacity evaluation of SM is presented in [10]. Under the assumption that the transmitted symbol s_k is selected as an i.i.d. Gaussian RV, [10] expresses the SM capacity as the sum of two parts

$$C_{\text{SM}} = W\left(C_1 + C_2\right), \tag{3}$$

where W is the available bandwidth. In (3), C_1 denotes the capacity that is linked to the standard symbol in the complex signal space and partially to the spatial symbol and is equal to

$$C_1 = \frac{1}{N_t} \sum_{i=1}^{N_t} \log_2 \left(1 + \frac{P\|\mathbf{h}_i\|_2^2}{N}\right), \tag{4}$$

where P denotes the average transmission power, $\mathbf{h}_i$ is the i-th column of the channel matrix $\mathbf{H} = [\mathbf{h}_1, \ldots, \mathbf{h}_{N_r}]$, and N is the noise power. Additionally, C_2 in (3) exclusively models the reminder of the spatial symbol part of SM, i.e., transmitting from a single antenna, and is equal to

$$C_2 = \frac{1}{N_t} \sum_{i=1}^{N_t} \int_y p\left(y|\mathbf{h}_i\right) \log_2 \frac{p\left(y|\mathbf{h}_i\right)}{p(y)} dy, \tag{5}$$

where $p\left(y|\mathbf{h}_i\right) = \frac{1}{\pi(P\|\mathbf{h}_i\|_2^2 + N)} \exp\left(-\frac{|y|^2}{P\|\mathbf{h}_i\|_2^2 + N}\right)$ represents the probability distribution function (pdf) of the received signal with respect to all receive antennas, and $p\left(y\right) = \frac{1}{N_t} \sum_{i=1}^{Nt} p\left(y|\mathbf{h}_i\right)$ denotes the pdf of the average received signal.

III. Power Consumption of a Base Station Employing Spatial Modulation

The EARTH power model is a very simple and elegant model that relates the transmitted power of a BS to the total power consumed [9]. Taking into account that some elements of a BS have a traffic load depended behavior and some others have a constant power consumption, the EARTH power model expresses the total power consumption of a BS as an affine function of the RF transmit power

$$P_{\text{supply}} = \begin{cases} N_{\text{RF}} P_0 + m N_{\text{RF}} P_{\text{t}}, & 0 < P_{\text{t}} \leq P_{\text{max}} \\ P_{\text{sleep}}, & P_{\text{t}} = 0. \end{cases} \tag{6}$$

In (6), P_{supply} denotes the total power supplied to the BS, N_{RF} denotes the number of RF chains at the BS, P_0 is the minimum power consumption per RF chain when the BS is active, m represents the slope of the load dependent power consumption, P_{t} is the RF transmit power per antenna, and P_{max} is the maximum transmitted power of an antenna. Finally, P_{sleep} is the power consumption when the sleep mode is applied. We note that the second part of (6), $mN_{\text{RF}}P_{\text{t}}$, corresponds to the total RF power transmission.

A brief inspection of (6) shows that the number of RF chains N_{RF} affects P_{supply} in two ways. The employment of more RF chains results in: (a) the increase of the load independent power consumption of $N_{\text{RF}}P_0$ and (b) the increase of the RF transmit power with a consequential increase of P_{supply}.

In order to quantify the power consumption gain of a multi-antenna BS employing SM ($N_{\text{RF}} = 1$) compared to a BS employing CMIMO ($N_{\text{RF}} = N_t$), we employ the EARTH power model of (6). For the case of the same total transmitted power between SM and CMIMO, using (6), it can be shown after some arithmetic manipulations that

$$P_{\text{supply}}^{\text{SM}} = P_{\text{supply}}^{\text{CMIMO}} - \left(N_{\text{RF}} - 1\right) P_0. \tag{7}$$

While, for the case where the power transmitted from every antenna in CMIMO is equal to the power transmitted in SM, it can be shown that

$$P_{\text{supply}}^{\text{SM}} = \frac{P_{\text{supply}}^{\text{CMIMO}}}{N_{\text{RF}}}. \tag{8}$$

In (7) and (8), $P_{\text{supply}}^{\text{SM}}$ denotes the total power supply of SM and $P_{\text{supply}}^{\text{CMIMO}}$ denotes the total power supply of a CMIMO scheme.

TABLE I
POWER MODEL PARAMETERS FOR DIFFERENT BS (SOTA 2010) [9].

BS type	P_0 (W)	m	P_{max} (W)	P_{sleep} (W)
Macro	118.7	2.66	40.0	63
Micro	53.0	3.1	6.3	-
Pico	6.8	4.0	0.13	-
Femtocell	4.8	7.5	0.05	-

In order to visualize the power saving of a BS employing SM compared to CMIMO, for the same transmit power and using the model parameters from Table I, in Fig. 1, we depict the RF transmit power versus P_{supply} for a macro BS with $N_t = 4$ transmit antennas. The model parameters of Table I represent the EARTH model for the year of 2010 [9]. In [9, 11], there is a future prediction for the EARTH power model in year 2014. In this work, we are mainly interested in presenting the EE of SM without taking into account any future hardware improvement. Thus, in the next sections we use Table I and refer to this case as State-of-the-Art 2010 (SotA 2010).

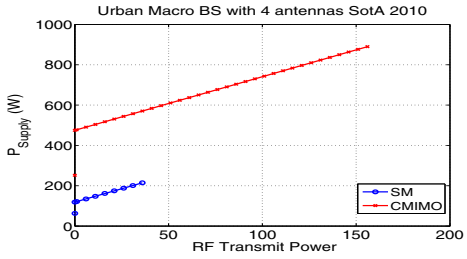


Fig. 1. RF transmit power versus P_{supply} for a macro BS with $N_t = 4$ transmit antennas (SotA 2010).

IV. RATE AND ENERGY EFFICIENCY RESULTS

In this Section, we use Monte Carlo simulations in order to evaluate the data rate and EE of different BSs (macro, micro, pico, and femtocell) in the downlink of a single cell. The studied schemes are SM, STBC [4], MISO with a single transmitting stream [1], and MIMO with N_t transmitting streams and channel knowledge only at the receiver [2]. Because we are interested in the limits, we use the Shannon capacity to get the average data rate and EE. For the case of SM, the capacity is given in (3), whereas for the benchmark techniques the capacity is given below. For the case of STBC, the capacity is written as

$$C_{\text{STBC}} = W\left[R\ \log_2\left(1+\frac{P}{NN_t}\|\mathbf{H}\|_F^2\right)\right], \qquad (9)$$

where R is the rate of a certain STBC code. For complex transmit signals, the rate of STBC is $R = 1$ only for the special case of $N_t = 2$. Generally, $R = \frac{1}{2}$ for $N_t > 2$ with an exception of three and and four transmit antennas, where $R = \frac{3}{4}$ [4]. Furthermore, the capacity of a MIMO system with N_t transmitting streams and channel state information only at the receiver is bounded by [2, eq. 8.17],

$$C_{\text{MIMO}} \leq W\left[m\log_2\left(1+\frac{P}{mNN_t}\|\mathbf{H}\|_F^2\right)\right], \qquad (10)$$

where $m = \min(N_t, N_r)$ and $\|\mathbf{H}\|_F^2 = \sum_{i=1}^{\min(N_t,N_r)}\lambda_i^2$. λ_i are the singular values of $\mathbf{H}$. The capacity of a MISO system with a single transmitting stream is given in [1, 2] as,

$$C_{\text{MISO}} = W\log_2\left(1+\frac{P\|\mathbf{h}\|_2^2}{N}\right). \qquad (11)$$

We note that due to the numerical evaluation of the integrals of (5), we assume the following approximation $C_{\text{SM}} \approx C_1$. This assumption is valid for two reasons: firstly, as stated in [10] the dominant part of C_{SM} is C_1 and C_2 is only a small portion of SM capacity. Secondly, the presented graphs for SM (Fig. 2 and 3) serve as a lower bound on the data rate and EE performance.

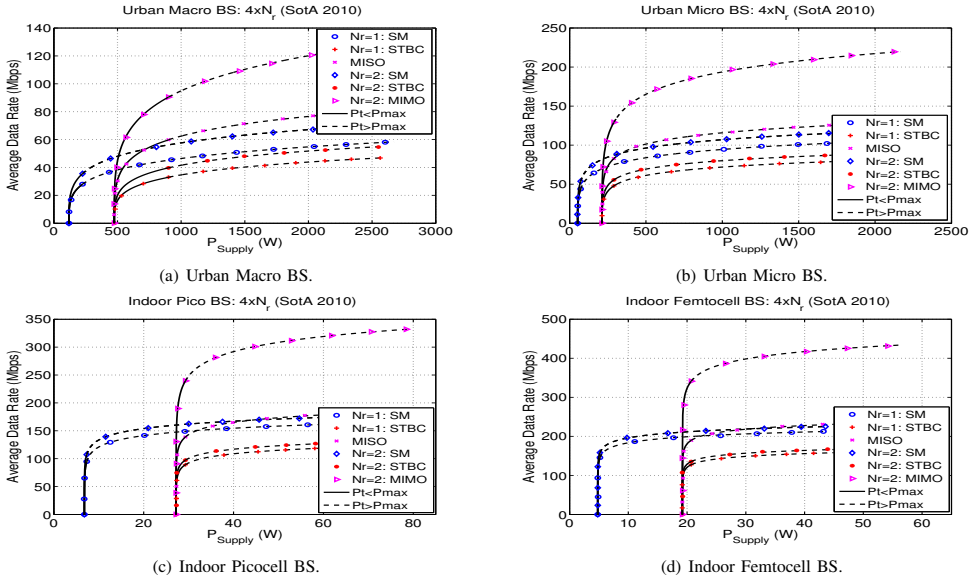
TABLE II
SIMULATION PARAMETERS.

Simulation Parameters	Values
BS Type	Macro, Micro, Pico, Femtocell
Power Model Parameters	SOTA 2010
Carrier Frequency	2 GHz
Path Loss Model	3GPP NLOS [12]
Iterations (Number of Channels)	100000
Bandwidth	10 MHz
Operating Temperature	Outdoor:290 K, Indoor: 293.5 K

TABLE III
LIMITS OF BS-USER DISTANCE.

BS type	d_{min} (m)	d_{max} (m)
Macro	10	5000
Micro	10	1000
Pico	10	150
Femtocell	5	50

Each type of BS is equipped with $N_t = 4$ antennas. Furthermore, we assume that the receiver is equipped with one or two antennas. In addition, we assume that the BS–to–user distance d is selected randomly following a spatially uniform distribution between d_{min} and d_{max}. Also, based on the type of BS, d_{min} and d_{max} are given in Table III. The shadowing standard deviation is 6 dB for the case of urban macro BS and 4 dB for the case of urban micro, indoor pico and femtocell BSs. In addition, the thermal noise power is defined as $N = W\kappa\theta$, where κ is the Boltzmann constant and θ is the operating temperature in K. Based on these assumptions, we conduct Monte Carlo simulations that quantify the average data rate and EE of a BS for a system configuration given in Table II. In order to calculate the average data rate of a BS, we follow an iterative procedure. At the i-th iteration, we select the distance d as described before and we compute the data rate $R_i(P_{\text{t}}^{\text{total}})$, where $P_{\text{t}}^{\text{total}} = N_{RF}P_t$. Finally, the average data rate is written as $\bar{R}(P_{\text{t}}^{\text{total}}) = \text{E}[\text{R}_\text{i}(\text{P}_\text{t}^{\text{total}})]$, where $\text{E}[\cdot]$ is the arithmetic mean operator.



(a) Urban Macro BS.

(b) Urban Micro BS.

(c) Indoor Picocell BS.

(d) Indoor Femtocell BS.

Fig. 2. Data Rates for different types of BS ($N_t = 4$). The solid lines depict the data rate achieved when the transmission power is restricted by the EARTH power model ($P_{\text{t}} \leq P_{\text{max}}$) and the dashed lines depict the data rate achieved when this restriction is not assumed, i.e., $P_{\text{t}} > P_{\text{max}}$.

A. Average Data Rate Results

Fig. 2 depicts the average data rate versus the power supply, when the BS type is urban macro, urban micro, indoor pico, and indoor femtocell. The depicted results correspond to two different scenarios of RF power transmission. In the first scenario (solid lines), the RF transmit power is restricted by the EARTH power model ($P_{\text{t}} \leq P_{\text{max}}$), which is a real-world hard constraint. While, in the second scenario (dashed lines), the RF transmit power is increased without any limitation.

Clearly, for $P_{\text{t}} \leq P_{\text{max}}$ (solid lines), in all cases there is a range of rates, where for the same data rate SM requires significantly less power supply compared to MISO and MIMO (a reduction which can reach up to 76%). Though, it can be seen that MISO and MIMO are capable of achieving higher data rates than SM with the significant cost of increasing the already high BS power consumption. As regard to the comparison of SM with STBC, SM achieves higher data rates with significantly less power consumption due to the code dependent fractional rates of STBC.

When the RF transmission power is increased beyond the EARTH limit of P_{max} (dashed lines), it can be seen from Fig. 2 that there is a crossing point where MISO and MIMO achieve higher data rates with less power consumption. This point happens when the RF transmission power is the dominant consuming factor in the power supply of a BS. For $N_r = 1$, this point occurs at 541, 249, 32, and 23 W for a macro, micro, pico, and femtocell BS, respectively, and for $N_r = 2$, this point occurs at 508, 223, 27, 19 W for a macro, micro, pico, and femtocell BS, respectively. However, these crossing points do not have a practical relevance, since P_{max} is a real-world hard constraint.

B. Energy Efficiency Results

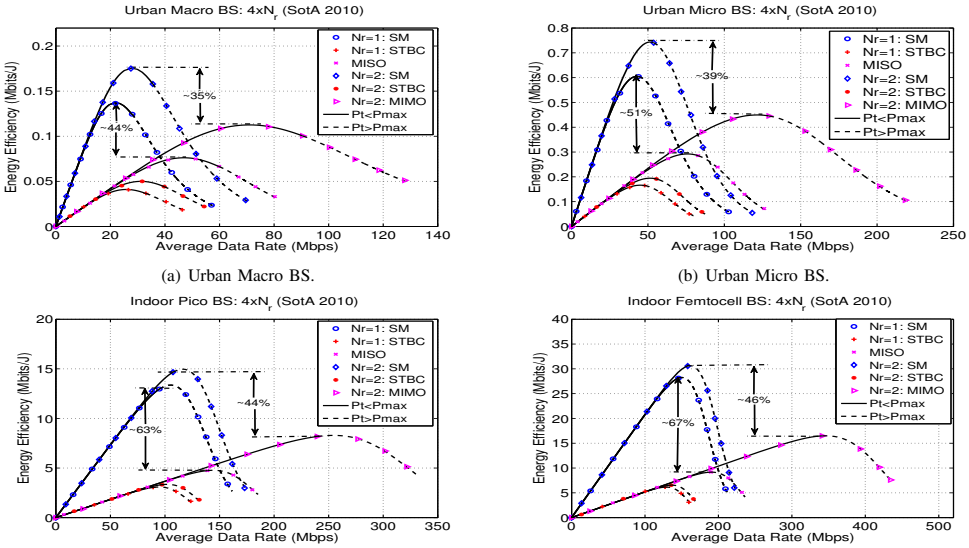
In this work, we define the EE of a BS, in terms of total power consumption, as

$$\text{EE}(P_{\text{t}}^{\text{total}}) = \frac{\bar{R}(P_{\text{t}}^{\text{total}})}{P_{\text{supply}}(P_{\text{t}}^{\text{total}})}. \tag{12}$$

The EE performance of different BS types, measured in Mbits/J, under different MIMO transmission schemes is presented in Fig. 3. As it can be seen from Fig. 3, SM is the most energy efficient transmission scheme for all types of BS, when $P_{\text{t}} \leq P_{\text{max}}$ (solid lines). This finding stems from the fact that SM has a lower constant power consumption due to the single RF chain configuration. An inspection of Fig. 3 reveals that MISO is up to 67% less energy efficient compared to SM and MIMO is up to 46% less energy efficient compared to SM. The EE difference between SM and STBC is even higher.

In addition, Fig. 3 shows that the EE of the studied schemes is maximized for a certain average data rate and power transmission point. The derivation of the point where EE is maximized is subject of our future work.

Finally, Fig. 3 shows that when the RF transmission power is increased above $P_{\text{t}} > P_{\text{max}}$ (dashed lines), there is a crossing point where MISO and MIMO become more energy efficient than SM. When the mobile terminal is equipped with one receive antenna, this crossing point occurs at 39.18, 72.5, 150.5, and 201.1 Mbps for a macro, micro, pico, and femtocell BS, respectively. And when the mobile terminal is equipped with two receive antennas, this crossing point occurs at 48.25, 82.57, 160.3, and 210.9 Mbps for a macro, micro, pico, and



(a) Urban Macro BS.

(b) Urban Micro BS.

(c) Indoor Picocell BS.

(d) Indoor Femtocell BS.

Fig. 3. EE Results for different types of BS ($N_t = 4$). The solid lines depict the EE achieved when the transmission power is restricted by the EARTH power model ($P_t \leq P_{\max}$) and the dashed lines depict the EE achieved when this restriction is not assumed, i.e., $P_t > P_{\max}$.

femtocell BS, respectively. As it is stated before, these crossing points do not have a practical value, since they violate the hard constraint of the maximum transmission power ($P_t \leq P_{\max}$).

As regard to the comparison between SM and STBC, we see that SM is always more energy efficient. Hence, taking into account that the current and future predicted hardware technology are restricted by the hard constraint of $P_{\max}$, SM is the most energy efficient scheme among the benchmark systems.

V. Conclusions

In this paper, using the limits of Shannon capacity and Monte Carlo simulations, we demonstrated that the single RF chain configuration of SM enables significant EE gains compared to MISO, MIMO and STBC (up to 67%), under all types of BSs. In more detail, we showed that for all evaluated multi-antenna schemes and types of BSs there is a transmission point where the EE is maximized. In addition, we concluded that for a range of data rates, SM offers the same data rate as the benchmark systems with significantly less power consumption, resulting into a reduction which can be as low as 76% for four transmit antennas.

Finally, we note that the main focus of our future work will be concentrated on the comparison of the EE of SM with different MIMO schemes and antenna configurations, combined with the development of an analytical framework for the simulation results.

Acknowledgement

This work has been funded by the Research Project GREENET (PITN-GA-2010-264759).

References

[1] G. J. Foschini and M. J. Gans, "On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas," *Wireless Personal Communications*, vol. 6, no. 6, pp. 311–335, 1998.

[2] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[3] P. Wolniansky, G. Foschini, G. Golden, and R. Valenzuela, "V-BLAST: an Architecture for Realizing very High Data Rates over the Rich-Scattering Wireless Channel," in *Unino Radio-Scientifique Internationale (URSI) Intern. Symp. on Signals, Systems, and Electronics (ISSSE)*, Sep. 29–Oct. 2, 1998, pp. 295–300.

[4] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time Block Codes from Orthogonal Designs," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1456–1467, Jul. 1999.

[5] A. Molisch and M. Win, "MIMO systems with antenna selection," *IEEE Microwave Magazine*, vol. 5, no. 1, pp. 46–56, Mar. 2004.

[6] L. Correia, D. Zeller, O. Blume, D. Ferling, A. Kangas, I. Godor, G. Auer, and L. Van Der Perre, "Challenges and Enabling Technologies for Energy Aware Mobile Radio Networks," *IEEE Communications Magazine*, vol. 48, no. 11, pp. 66–72, Nov. 2010.

[7] M. Di Renzo, H. Haas, and P. M. Grant, "Spatial Modulation for Multiple-Antenna Wireless Systems: A Survey," *IEEE Commun. Mag.*, vol. 49, no. 11, pp. 182–191, Nov. 2011.

[8] J. Jeganathan, A. Ghrayeb, and L. Szczecinski, "Spatial Modulation: Optimal Detection and Performance Analysis," *IEEE Commun. Lett.*, vol. 12, no. 8, pp. 545–547, 2008.

[9] G. Auer, V. Giannini, I. Godor, P. Skillermark, M. Olsson, M. Imran, D. Sabella, M. Gonzalez, C. Desset, and O. Blume, "Cellular Energy Efficiency Evaluation Framework," in *Proc. of the Vehicular Technology Conference (VTC Spring)*, May 2011, pp. 1–6.

[10] Y. Yang and B. Jiao, "Information-Guided Channel-Hopping for High Data Rate Wireless Communication," *IEEE Communications Letters*, vol. 12, no. 4, pp. 225 –227, Apr. 2008.

[11] H. Holtkamp, G. Auer, S. Bazzi, and H. Haas, "Minimizing Base Station Power Consumption," *IEEE Journal on Selected Areas in Communications*, vol. PP, no. 99, pp. 1–10, 2013.

[12] 3GPP, "Further Advancements for E-UTRA Physical Layer Aspects (Release 9)," 3GPP TR 36.814 V0.4.1 (2009-02), Sep. 2009. Retrieved Jun. 2, 2009 from www.3gpp.org/ftp/Specs/.

F.7 IEEE 18th Int. Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD) – 1st paper

2013 IEEE 18th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)

A Base Station Switching On-Off Algorithm Using Traditional MIMO and Spatial Modulation

Athanasios Stavridis[(1)], Sandeep Narayanan[(2,3)], Marco Di Renzo[(4)], Luis Alonso[(5)], Harald Haas[(1)], and Christos Verikoukis[(6)]
[(1)] Institute for Digital Communications (IDCOM), The University of Edinburgh, Edinburgh, UK
[(2)] WEST Aquila s.r.l., L'Aquila, Italy
[(3)] Center of Excellence for Research DEWS, University of L'Aquila, Italy
[(4)] Laboratory of Signals and Systems (L2S), UMR 8506 CNRS – SUPELEC – Univ Paris–Sud, Paris, France
[(5)] Signal Theory and Communications Dept., Technical University of Catalonia, Barcelona, Spain
[(6)] Telecommunications Technological Center of Catalonia, Barcelona, Spain
E–Mail: a.stavridis@ed.ac.uk, sandeep.narayanan@westaquila.com, marco.direnzo@lss.supelec.fr, luisg@tsc.upc.edu, cveri@cttc.es, h.haas@ed.ac.uk

***Abstract*—In this paper, we propose a novel rate adaptive Base Station (BS) switch on-off algorithm that employs Spatial Modulation (SM) and traditional Multiple-Input Multiple-Output (MIMO). We consider a cluster of BSs and based on the required average data rate, our algorithm selects the most energy efficient scheme between SM and MIMO with only Channel State Information at the Receiver (MIMO-CSIR). In order to further increase the Energy Efficiency (EE) (in bits/J), we activate or deactivate BSs in conjunction to switching between different transmission modes (SM or MIMO-CSIR). In both cases, we ensure that the User Terminals (UTs) do not face a reduction in Quality-of-Service (QoS) in the sense of ergodic capacity. Using the fundamental limit of Shannon capacity and a BS power model, we show that the use of SM, in such an algorithm, offers significant EE improvements. Finally, we demonstrate simulation results which show that, depending on the required average data rate, our algorithm achieves the same or better EE compared to the employed benchmark systems.**

I. Introduction

During the recent years, Energy Efficiency (EE) in wireless cellular networks has attracted considerable attention in research. This interest is mainly contributed to economical and environmental reasons. From an operator's point of view, an efficient way to reduce its operational cost is to decrease the energy consumption of the mobile networks. In addition, the increasing evidence for the man-made climate change strongly drives the research on more energy efficient (green) solutions in wireless communication.

The main energy consumer in a cellular network is the Base Station (BS) [1]. Hence, the intelligent use of the BSs is becoming increasingly important. In order to achieve EE in a cellular network the following strategies can be applied: i) advanced hardware utilization at the BSs [2], ii) selection of the most energy efficient physical layer technique for a given data load [3, 4], and iii) intelligent deployment of BSs by switching off a subset of them during the low data traffic periods [5]. In any case, it is very important that the chosen solution does not affect the Quality of Service (QoS) experienced by the users. In this paper, we follow a cross-strategy method having as an objective to jointly exploit the EE offered by the selection of the appropriate physical method and the deactivation of a subset of BSs.

Research on Multiple-Input Multiple-Output (MIMO) wireless communication demonstrates that multi-antenna communication is a promising solution for the spectrally efficient future wireless networks. The employment of multi-antenna elements, both at the transmitter and receiver, is shown to offer increased system capacity without the need for extra radio resources [6].

Unfortunately, multi-antenna elements face a major disadvantage of requiring multiple Radio Frequency (RF) chains. RF chains are expensive circuits that do not follow Moore's law [7]. Thus, the real system implementation becomes expensive and often impractical. In addition, RF chains are electronic circuits that may reduce the EE of a wireless system. RF chains include Power Amplifiers (PAs) which are responsible for 50-80% of the total power consumption in the transmitter [8]. Hence, there is a significant interest in new MIMO schemes that do not require multiple RF chains. An example of such a scheme is Spatial Modulation (SM) [9, 10].

The authors of [3, 4] demonstrate the EE improvements of SM at a BS compared to the traditional schemes of MIMO with only Channel State Information at the Receiver (MIMO-CSIR) [6] and Space-Time Block Coding (STBC) [11]. Furthermore, [3, 4] show that a BS employing SM requires less total energy compared to a BS with traditional MIMO transmission. Motivated from this fact, we aim to exploit the EE of SM in a cluster of BSs in conjunction with the BS switch on-off methodology. We note that we do not consider Channel State Information at the Transmitter (CSIT).

We propose a simple but yet very effective four stage algorithm which offers increased EE without affecting the QoS that users experience in the ergodic capacity sense. In each stage of our algorithm, the most energy efficient physical layer method (SM or MIMO-CSIR) is selected, in addition to the activation/deactivation of the appropriate number of BSs that are sufficient to offer the desired data rate. Using the fundamental limit of Shannon capacity and Monte Carlo simulations, we show that the use of SM and MIMO-CSIR is

978-1-4673-4919-2/13/$31.00 ©2013 IEEE

2013 IEEE 18th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)

more energy efficient than that of: i) Single-Input Multiple-Output (SIMO) and MIMO-CSIR and ii) only MIMO-CSIR, for the same BS switch on-off strategy.

The remainder of this paper is organized as follows: Section II presents the cellular system model that we consider. The power model of a BS under different MIMO transmission schemes is introduced in Section III. Section IV demonstrates the employed physical layer techniques and evaluates their achievable Shannon capacity with respect to the BS power supply. The proposed algorithm is introduced in Section V. Section VI provides the rate and EE results for the proposed scheme as well as the employed benchmark systems. Finally, we conclude this paper in Section VII.

Notation: In the following, lowercase bold letters denote vectors and uppercase bold letters denote matrices. Finally, $\|\cdot\|_2$ and $\|\cdot\|_F$ stand for the Euclidean and Frobenius norm, respectively.

II. Cellular System Model

The cellular system model that we consider is depicted in Fig. 1. In this model, an urban geographical area accommodates 7 BSs with inter-site distance d_{is} and cell radius d_{cr}. We concentrate on the downlink and assume that each BS is equipped with N_t antennas and each one of the K User Terminals (UTs) is equipped with N_r antennas. In order to form a multiuser system, we employ Time Division Multiple Access (TDMA), where in each time slot a UT is served by a certain BS that aims to provide the requested data rate. Although Inter-Cell Interference (ICI) is a limiting factor for the cellular networks, in this work, we assume that there is no ICI between different cells.

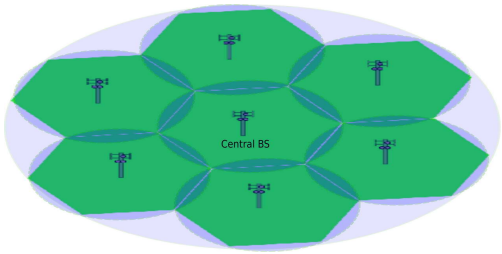


Fig. 1. The cellular system model.

We assume a seven-cell cellular set up as depicted in Fig. 1. In addition to the spectral efficiency and the QoS experienced by the uniformly distributed UTs in the network, we are also interested in the energy consumption and EE. An effective way to fulfill these requirements without sacrificing one of them is to switch on-off BSs in conjunction with intelligent adaptive MIMO. Hence, we assume that all BSs are connected with a processing unit which decides if a BS is active or not, depending on the network traffic. Furthermore, we assume that each BS is able to switch between different transmission modes by activating all of its RF chains or by keeping active only one RF chain. When all the RF chains of a BS are active, the transmission mode is MIMO-CSIR. While, when only one RF chain is active the transmission mode is either SM or SIMO. In Section VI, using Monte Carlo simulations, we compare the EE results between a cellular system which employs SM and MIMO-CSIR with: i) a system which employs SIMO and MIMO-CSIR, and ii) a system which employs only MIMO-CSIR.

III. Power Model of a Multi-Antenna Base Station

In order to quantify the power (energy) consumption and the EE of a cellular system, we focus on the power supply of a single BS and then scale appropriately to the whole network. A BS is composed of elements like multiple transceivers, PAs, baseband interface, Alternating Current-Direct Current (AC-DC) converter, DC-DC power supply, and a cooling system. Each of these elements has a particular power consumption which is either load or non-load dependent. A simple model, yet very insightful, that connects the transmission power of a BS with its total power supply is the Energy Aware Radio and neTwork tecHnology (EARTH) power model [2]. It is shown in [2] that the power supply of a BS is approximated by a linear function of the RF transmit power. In more detail, the EARTH power model expresses the BS power supply as:

$$P_{\text{s}} = \begin{cases} N_{\text{RF}}P_0 + mN_{\text{RF}}P_{\text{t}}, & 0 < P_{\text{t}} \leq P_{\text{max}} \\ P_{\text{sleep}}, & P_{\text{t}} = 0, \end{cases} \quad (1)$$

where P_{s} is the BS power supply, P_0 is the BS power consumption when the BS is active and transmits at its minimum RF power, N_{RF} is the number of RF chains employed by the BS, m is the slope of the load dependent power consumption, P_t is the RF transmit power, and P_{sleep} is the power consumption when the BS is in the sleep mode.

A brief inspection of (1) shows that the number of RF chains (N_{RF}) affects P_{s} in two ways. The use of more RF chains results in: (a) the increase of the load independent power described as $N_{\text{RF}}P_0$ and (b) the increase of the total RF transmit power, *i.e.*, an increase of P_{s}. In [3], it is shown that for the same RF transmit power, the total power saving of a single RF scheme (like SM) scales linearly with the number of RF chains required by traditional multi-RF chain MIMO schemes.

IV. Physical Layer Transmission Techniques

In this section, we present three physical layer transmission techniques that can be employed by the BSs of Fig. 1. In addition, we demonstrate their fundamental limit of Shannon capacity with respect to the BS power supply. The physical layer techniques that we consider are: i) MIMO-CSIR [6], ii) SIMO [6], and iii) SM [9]. SM and SIMO operate with a single RF chain at the transmitter. Hence, the BS power supply during the activation period is given from (1) as

$$P_s = P_0 + mP_{\text{t}}. \quad (2)$$

In contrast, MIMO-CSIR requires multiple RF chains at the transmitter, resulting in BS power supply equal to

$$P_s = N_{\text{RF}}P_0 + mP'_t, \quad (3)$$

where $P'_t = N_{\text{RF}}P_{\text{t}}$. In this paper, we are interested in open loop schemes, thus we do not consider CSIT.

2013 IEEE 18th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)

MIMO-CSIR forms N_t parallel streams from the transmitter to the receiver. Due to the lack of CSIT, the transmitted power is equally distributed to N_t parallel streams. The Shannon capacity for MIMO-CSIR is given in [12, eq. 8.17]. If we incorporate the power model of (1), the capacity that can be achieved with respect to the BS power supply is bounded as

$$C_{\text{MIMO}}(P_s) \leq nW \log_2 \left(1 + \frac{P_s - N_t P_0}{nmN_t N} \|\mathbf{H}\|_F^2\right), \quad (4)$$

where $n = \min(N_t, Nr)$ and

$$\|\mathbf{H}\|_F^2 = \sum_{i=1}^{n} \lambda_i^2.$$

In (4), λ_i, $i = 1, \ldots, n$ denote the singular values of $\mathbf{H}$, W represents the transmission bandwidth and N is the noise power.

An alternative scheme that can be employed at the BS is SIMO. In SIMO, a single symbol stream is established from the transmitter to the receiver. As the transmitter activates a single antenna during the signaling period, SIMO operates with a single RF chain. In this case, the achievable Shannon capacity with respect to the BS power supply is

$$C_{\text{SIMO}}(P_s) = W \log_2 \left(1 + \frac{(P_s - P_0)\|\mathbf{h}\|_2^2}{mN}\right), \quad (5)$$

where $\mathbf{h}$ is the SIMO channel.

SM is a MIMO scheme that enjoys a single RF chain configuration at the transmitter and at the same time is able to achieve a spatial multiplexing gain [9, 10]. Its transmission mechanism operates as follows. The bit stream to be transmitted during a signaling period is divided into two sequences. The first sequence is encoded into the signal domain using a signal symbol drawn from a conventional constellation like Quadrature Amplitude Modulation (QAM). While, the second sequence is encoded using the index of the single activated antenna. At the receiver side, an optimal ML detector is employed in order to jointly detect the transmitted signal and the activated antenna. Because of its transmission principle, SM does not require Inter-Antenna Synchronization (IAS) and totally avoids inter-channel interference at the receiver.

The information theoretic analysis of SM is presented in [13]. If we incorporate the power model of (2), under the assumption that the transmitted symbol is selected as an i.i.d. Gaussian Random Variable (RV), the Shannon capacity of SM is expressed as the sum of two parts:

$$C_{\text{SM}}(P_s) = W\left(C_1(P_s) + C_2(P_s)\right). \quad (6)$$

The first component of (6) denotes the capacity that is linked to the standard symbol in the complex signal space and partially to the spatial symbol and is equal to

$$C_1(P_s) = \frac{1}{N_t} \sum_{i=1}^{N_t} \log_2 \left(1 + \frac{(P_s - P_0)\|\mathbf{h}_i\|_2^2}{mN}\right), \quad (7)$$

where $\mathbf{h}_i$ is the i-th column of the MIMO channel matrix $\mathbf{H} = [\mathbf{h}_1, \ldots, \mathbf{h}_{N_r}]$ between the serving BS and the receiving UT. In addition, the second part of (6) exclusively models the reminder of the spatial symbol part of SM, *i.e.*, transmitting from a single antenna, and is equal to

$$C_2(P_s) = \frac{1}{N_t} \sum_{i=1}^{N_t} \int_y p\left(y|\mathbf{h}_i\right) \log_2 \frac{p\left(y|\mathbf{h}_i\right)}{p(y)} dy, \quad (8)$$

where,

$$p\left(y|\mathbf{h}_i\right) = \frac{1}{\pi\left(\frac{P_s - P_0}{m}\|\mathbf{h}_i\|_2^2 + N\right)} \exp\left(-\frac{|y|^2}{\frac{P_s - P_0}{m}\|\mathbf{h}_i\|_2^2 + N}\right)$$

represents the probability distribution function (pdf) of the received signal with respect to all receive antennas. While,

$$p\left(y\right) = \frac{1}{N_t} \sum_{i=1}^{N_t} p\left(y|\mathbf{h}_i\right)$$

denotes the pdf of the average received signal. As we observe from (8), $C_2(P_s)$ requires the difficult numerical evaluation of a summation of integrals. Thus, we choose to approximate $C_2(P_s)$ with zero in low Signal-to-Noise Ratio (SNR). While, for high SNR, we use the fact that $C_2(P_s) \approx \log_2(N_t)$ [14].

V. Proposed Algorithm

The authors of [3, 4] study the EE of SM under different types of BS (macro, micro, pico, and femtocell) and different antenna configurations. These papers demonstrate that under every type of BS and every antenna configuration: i) SM requires less power supply than any other multi-RF chain architecture in order to transmit the same amount of RF power, ii) SM achieves a region of rates in $(0, R_{\text{max}}]$, where R_{max} is the maximum achievable rate given a real-world hard constraint on the transmit power, with significantly less power supply, iii) the EE of SM for rates within $(0, R_{\text{max}}]$ is significantly higher than any other traditional open loop MIMO scheme, and iv) traditional MIMO schemes offer rates higher than SM by increasing the already high power consumption.

Motivated from the previous conclusions in a single BS, we aim to extent the EE gains of SM to the cellular framework illustrated in Fig. 1. We propose a rate adaptive algorithm which switches on-off BSs in addition to switching between SM and traditional MIMO. In more detail, our algorithm operates in 4 stages:

- **Stage 1:** when the required data rate is low, only the central BS is activated using SM.
- **Stage 2:** when the required data rate is increased and cannot be achieved by SM, the transmission scheme of the central BS is turned to MIMO-CSIR.
- **Stage 3:** when the required data rate is further increased and cannot be achieved by the central BS, all the BSs are activated using SM.
- **Stage 4:** when the required data rate is further increased, all the BSs are active using MIMO-CSIR.

The main motivation behind Stage 1 is that the geographical area that needs to be covered by the cellular network is served by the central BS using the energy efficient scheme of SM. During Stage 2, the cellular network employs MIMO-CSIR at the central BS, but still achieves reduced power (energy)

2013 IEEE 18th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)

consumption by keeping all the other BSs deactivated. When the central BS is unable to achieve the desired data rate, all the BSs are activated using SM. Although all the BSs are active during Stage 3, high EE is maintained due to the use of SM at each BS. Finally, when the data rate requirement is at its peak, all the BSs switch their transmission mode to MIMO-CSIR.

VI. Rate and Energy Efficiency Results

In this Section, we employ Monte Carlo simulations in order to evaluate the power consumption and the EE of the adaptive MIMO BS switch on-off algorithm of Section V. We assume that each BS is equipped with $N_t = 4$ transmit antennas. When all BSs are active (Stage 3 and 4), each BS covers a circular area of radius of 800 m. While, when only the central BS is active (Stage 1 and 2), this BS covers a circular area of radius of 2400 m. In addition, we assume that the cluster of BSs serves multiple UTs that are randomly distributed (using a uniform distribution) inside the region that needs to be covered (a circular area of radius of 2400 m). We note that we exclude the area that has a distance less than 10 m from a BS. We follow this assumption because the urban Macro BSs are usually installed on top of buildings. Each UT is equipped with a single antenna. Multiuser access is achieved by assuming TDMA, where each BS serves a single UT at each time slot (one per multiuser frame). Every channel realization between a BS and a UT is generated using the path loss model which is presented in Table I. The thermal noise power is defined as $N = W\kappa\theta$, where W is the available bandwidth, κ is the Boltzmann constant and θ is the operating temperature in K. Based on these assumptions and using the simulation parameters of Table I, we conduct Monte Carlo simulations that quantify the average data rate and EE of the cellular system of Fig. 1 when the algorithm of Section V is employed. For the EARTH power model of (1), we employ the following values: $P_0 = 118.7$ W, $P_{\max} = 40$ W, $P_{\text{sleep}} = 63$ W, and $m = 2.66$, which are given in [15] and correspond to a State-of-the-Art (SotA) urban Macro BS in year 2010 [16].

Because we are interested in the fundamental limits of the network data rate, we use the Shannon capacity in order to get the average data rate and EE. The Shannon capacity for each employed physical layer methods is given in (4), (5), and (6). In this paper, we compute the average sum rate as

$$\bar{R}_{\text{total}} = \sum_{i=1}^{K} \bar{R}_i,$$

where $\bar{R}_i$ is the average reception data rate achieved by the i-th UT. Finally, the EE (bits/J) is expressed as

$$\text{EE}(P_{\text{t}}^{\text{total}}) = \frac{\bar{R}_{\text{total}}}{P_{\text{s}}^{\text{total}}}, \quad (9)$$

where $P_{\text{s}}^{\text{total}}$ is the power supply of the cellular network and depends on the number of active BSs. Finally, we note that we do not take into account the power supply for the wired backhaul communication between different BSs, which is only a small portion of the sum power supply of the BSs [1].

In order to understand the results of Fig. 3 and 4, in Fig. 2 we present the EE results for a single BS when: i) the cell radius is 2400 m, which corresponds to the case where only the central BS is active and ii) the cell radius is 800 m, which is the case of a single cell when all BSs are active. As we see from Fig. 2, in both cases, SM offers the highest EE among the demonstrated transmission schemes. In addition, we observe that as the cell radius is decreased, the maximum EE is rapidly increased due to the less severe effect of the channel path loss. Hence, we expect that the algorithm of Section V will outperform the EE performance of every other benchmark scheme that does not employ SM.

TABLE I
Simulation Parameters.

Simulation Parameters	Values
BS Type	Urban Macro BS
d_{is}	$800\sqrt{3}$ m
d_{cr}	800 m
Carrier Frequency	2 GHz
Path Loss Model	3GPP NLOS [17]
Shadowing Standard Deviation	6 dB
Iterations (Number of Channels)	100000
Bandwidth	10 MHz
Operating Temperature	Outdoor:290 K

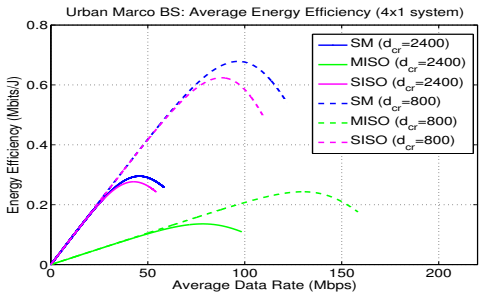


Fig. 2. Average data rate versus EE for an urban Macro BS with different cell radius d_{cr}.

In Fig. 3 and 4, we present the cellular power consumption versus the average sum rate $\bar{R}_{\text{total}}$ and the average sum rate versus the cellular EE, respectively, for the algorithm of Section V. To compare, in the same figures, we present the performance of two benchmark systems. Both benchmark systems operate similarly as the algorithm in Section V, but do not employ SM. The first benchmark system, which is termed as Benchmark 1, employs only MISO-CSIR during the activation of the central BS and the activation of all the BSs. The second benchmark system, which is denoted as Benchmark 2, operates the same as our algorithm, but instead of employing SM during stages 1 and 3, employs SISO transmission.

In Fig. 3, we observe a number of discontinuities in the resulting graphs. These discontinuities correspond to the points of transition from one active BS to seven active BSs. At these points, the cellular power consumption is rapidly increased. Furthermore, an inspection of Fig. 3 shows that our algorithm is able to offer the same data rate as the benchmark systems

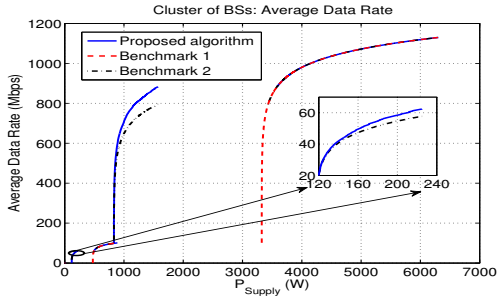


Fig. 3. Power supply versus average data rate for the cellular system of Fig. 1.

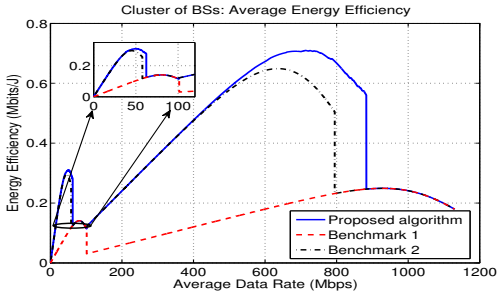


Fig. 4. Average data rate versus EE for the cellular system of Fig. 1.

with less or at most with the same supply power. The proposed algorithm significantly reduces the power consumption of the cellular network when the transmission mode is SM. When the transmission mode is changed to MISO-CSIR, the proposed algorithm consumes the same supply power as the benchmark systems.

Fig. 4 presents the EE results for the new algorithm and the used benchmark systems. Clearly, the new algorithm achieves higher EE, for a region of data rates. This is because the algorithm benefits from the single RF transmission in SM. In addition, we observe that when the proposed algorithm is applied, the EE of the cellular system is maximized at two points: a) at about 50 Mbps, when only the central BS is active and b) at about 750 Mbps when all the BSs are active. This observation highlights that advanced scheduling may be applied among different BSs in order to further increase the system EE. Such a study is out of scope here and will be considered in future work. Finally, we observe that the new algorithm offers the same EE as the benchmark systems, during Stage 2 and 4, when MISO-CSIR is employed.

VII. Conclusions

In this paper, we investigated the problem of switching on-off BSs in conjunction with the selection of the most EE transmission mode (SM or MIMO-CSIR), given a certain target data rate. By considering a cluster of BSs and using the fundamental limit of Shannon capacity and a BS power model, we concluded that the proposed algorithm overcomes or offers the same performance as the employed benchmark systems in terms of EE (bits/J) as well as decreased power consumption (W). Special care was given in order to ensure that the QoS (in mean sense) that the UTs experience was not decreased. Finally, we demonstrated that the use of SM was instrumental in achieving the significant EE improvements.

Acknowledgement

This work has been funded by the Research Project GREENET (PITN-GA-2010-264759).

References

[1] C. Han, T. Harrold, S. Armour, I. Krikidis, S. Videv, P. Grant, H. Haas, J. Thompson, I. Ku, C.-X. Wang, T. A. Le, M. Nakhai, J. Zhang, and L. Hanzo, "Green Radio: Radio Techniques to Enable Energy-efficient Wireless Networks," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 46–54, Jun. 2011.

[2] G. Auer, V. Giannini, C. Desset, I. Godor, P. Skillermark, M. Olsson, M. Imran, D. Sabella, M. Gonzalez, O. Blume, and A. Fehske, "How Much Energy is Needed to Run a Wireless Network?" *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 40–49, 2011.

[3] A. Stavridis, S. Sinanović, M. D. Renzo., H. Haas, and P. Grant, "An Energy Saving Base Station Employing Spatial Modulation," in *IEEE 17th Int. Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, Sep. 17–19 2012, pp. 231 –235.

[4] A. Stavridis, S. Sinanović, M. D. Renzo., and H. Haas, "Energy Evaluation of Spatial Modulation at a Multi-Antenna Base Station," in *Proc. of the 78th IEEE Veh. Tech. Conf. (VTC)*, Las Vegas, USA, Sep. 2–5, 2013.

[5] A. Bousia, A. Antonopoulos, L. Alonso, and C. Verikoukis, "Green Distance-Aware Base Station Sleeping Algorithm in LTE-Advanced," in *Proc. of the 2012 IEEE International Conference on Communications (ICC)*, 2012, pp. 1347–1351.

[6] G. J. Foschini and M. J. Gans, "On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas," *Wireless Personal Communications*, vol. 6, no. 6, pp. 311–335, 1998.

[7] A. Molisch and M. Win, "MIMO systems with antenna selection," *IEEE Microwave Magazine*, vol. 5, no. 1, pp. 46–56, Mar. 2004.

[8] L. Correia, D. Zeller, O. Blume, D. Ferling, A. Kangas, I. Godor, G. Auer, and L. Van Der Perre, "Challenges and Enabling Technologies for Energy Aware Mobile Radio Networks," *IEEE Communications Magazine*, vol. 48, no. 11, pp. 66–72, Nov. 2010.

[9] M. Di Renzo, H. Haas, and P. M. Grant, "Spatial Modulation for Multiple-Antenna Wireless Systems: A Survey," *IEEE Commun. Mag.*, vol. 49, no. 11, pp. 182–191, Nov. 2011.

[10] M. D. Renzo, H. Haas, A. Ghrayeb, S. Sugiura, and L. Hanzo, "Spatial Modulation for Generalized MIMO: Challenges, Opportunities and Implementation," 2013. [Online]. Available: http://hal-supelec.archives-ouvertes.fr/hal-00840278

[11] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time Block Codes from Orthogonal Designs," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1456–1467, Jul. 1999.

[12] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[13] Y. Yang and B. Jiao, "Information-Guided Channel-Hopping for High Data Rate Wireless Communication," *IEEE Communications Letters*, vol. 12, no. 4, pp. 225 –227, Apr. 2008.

[14] J. Jeganathan, A. Ghrayeb, L. Szczecinski, and A. Ceron, "Space Shift Keying Modulation for MIMO Channels," *IEEE Trans. on Wireless Commun.*, vol. 8, no. 7, pp. 3692–3703, Jul. 2009.

[15] G. Auer, V. Giannini, I. Godor, P. Skillermark, M. Olsson, M. Imran, D. Sabella, M. Gonzalez, C. Desset, and O. Blume, "Cellular Energy Efficiency Evaluation Framework," in *Proc. of the Vehicular Technology Conference (VTC Spring)*, May 2011, pp. 1–6.

[16] H. Holtkamp, G. Auer, S. Bazzi, and H. Haas, "Minimizing Base Station Power Consumption," *IEEE Journal on Selected Areas in Communications*, vol. PP, no. 99, pp. 1–10, 2013.

[17] 3GPP, "Further Advancements for E-UTRA Physical Layer Aspects (Release 9)," 3GPP TR 36.814 V0.4.1 (2009-02), Sep. 2009. Retrieved Jun. 2, 2009 from www.3gpp.org/ftp/Specs/.

F.8 IEEE 18th Int. Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD) – 2nd paper

2013 IEEE 18th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)

Distributed Spatially-Modulated Space-Time-Block-Codes

Sandeep Narayanan[1,2], Athanasios Stavridis[3], Marco Di Renzo[4], Fabio Graziosi[2], Harald Haas[3]
[1] WEST Aquila s.r.l., Via G. Gronchi 18, Nucleo Industriale di Pile, 67100 L'Aquila, Italy
[2] University of L'Aquila, Center of Excellence for Research DEWS, Via G. Gronchi 18, Nucleo Industriale di Pile, 67100 L'Aquila, Italy
[3] The University of Edinburgh, Institute for Digital Communications (IDCOM), Mayfield Road, Edinburgh, EH9 3JL, UK
[4] Laboratory of Signals and Systems (L2S), UMR 8506 CNRS – SUPELEC – Univ Paris–Sud, 3 rue Joliot–Curie, 91192 Gif–sur–Yvette (Paris), France
E–Mail: sandeep.narayanan@westaquila.com, a.stavridis@ed.ac.uk, marco.direnzo@lss.supelec.fr, fabio.graziosi@univaq.it, h.haas@ed.ac.uk

***Abstract*—In this paper, a novel energy–efficient protocol, intended for wireless networks with large number of relay nodes, is proposed. The main distinguishable feature of the proposed protocol is that it offers throughput enhancement, by having the same diversity–gain and having to activate the same number of relay nodes at any given time–instance, as the the conventional distributed space–time–block–codes (D-STBC). The proposed protocol applies the recently introduced idea of Spatially-Modulated Space–Time–Block–Codes (SM-STBC) devised for multiple–input–multiple–output (MIMO) systems to cooperative relay networks. The specific contributions of this paper are: i) an error–aware Maximum–Likelihood (ML) demodulator, which is robust to demodulation errors at the relays is developed; ii) a low–complexity error–aware demodulator is also developed; and iii) it is shown, with the help of Monte Carlo simulations, that the proposed protocol outperforms state–of–the–art relaying protocols.**

I. Introduction

It is well known that multiple–input–multiple–output (MIMO) technology can significantly enhance the performance of communication systems [1]–[3]. This performance enhancement is achieved by utilizing co-located multiple antennas at the transmitter and/or receiver. Unfortunately, there are practical constraints in equipping the mobile terminals with multiple antennas. However, relaying and distributed cooperation can mimic the performance advantages of MIMO systems by forming a virtual antenna array using the single-antenna cooperating nodes [4].

Distributed space–time–block–coded (D-STBC) cooperative protocols have been studied extensively in the literature in the recent years [5], [6], [8]. In [5], Laneman and Wornell showed that space-time coded cooperative diversity protocols can offer full spatial-diversity as the co–located MIMO. Thereafter, many attempts have been made to incorporate space–time–block–codes (STBC) with regenerative [6]–[8] and non-regenerative relay networks [9]. In [6] and [7], a Maximum–Likelihood (ML) demodulator, which takes advantage of the knowledge of the instantaneous error–probability of source–relay link, is employed at the destination. These kind of receivers are robust to demodulation errors at the relays and are generally termed as error-aware demodulators. In [8], perfect demodulation at the relays is achieved via Cyclic Redundancy Check (CRC) operations.

Recently, Spatial Modulation (SM) has been proposed as a low–complexity MIMO technique [10]–[12]. SM exploits spatial dimension in addition to signal constellation to convey the information bits. Although SM is originally proposed as an Inter-Channel Interference (ICI) free spatial–multiplexing technique, many attempts have been made recently to design transmit–diversity achieving SM scheme [14]–[17]. The application of SM to relay networks have been considered in [18]–[23]. However, no attempts have been made to apply the transmit–diversity achieving SM schemes to relay networks.

Against this background, in this paper, we propose a novel protocol for wireless networks with potentially large number of single–antenna relay nodes. The proposed protocol exploits the concept of transmit–diversity achieving SM scheme proposed in [14] for co–located MIMO systems. This scheme effectively combines the features of SM and STBC to achieve transmit–diversity gain.

The main distinguishable feature of the protocol proposed in this paper is that, it offers throughput enhancement, by having the same diversity–gain and having to activate the same number of relay nodes at any given time–instance, as the the conventional D-STBC scheme [7]. Alamouti code is chosen as the core STBC through out this paper. The extension to more general STBC schemes is possible, but it is not considered in this paper. Conventional distributed Alamouti schemes provides a normalized rate of $\frac{1}{2}\log_2(M)$ bits/CP [7], where M is the constellation size of the transmitted signal from the source and CP stands for cooperation–phase. On the other hand, using the scheme proposed in this paper, a normalised rate of $\frac{2}{5}\log_2(M) + \frac{1}{5}\log_2(\lfloor N_r/2 \rfloor)$ bits/CP, can be achieved, where N_r and $\lfloor . \rfloor$ represents the number of relay nodes and the floor integer, respectively. Clearly, the rate is much higher when there are large number of cooperating relay nodes. Moreover, the increased rate is achieved by activating only two of the N_r available relay at any given time. Hence, by using the proposed scheme, higher throughout can be achieved using the same total average transmit energy at the relays as the distributed Alamouti scheme [7]. The specific contributions of this paper are as follows: i) a novel energy-efficient protocol for wireless networks with large number of relay nodes, is proposed; ii) an error–aware ML demodulator at the destination, which is robust to demodulation errors at the relays is developed; iii) a low-complexity demodulator, which simplifies the structure of the ML demodulator is developed; and iv) the performance in terms of Average–Symbol–Error–Probability and energy gain in terms of Relative–Average–Energy–Reduction (RAER) per bit is studied. Furthermore, with the help of Monte-Carlo simulations it is shown that the proposed scheme outperforms other state–of–the–art relaying protocols.

The rest of the paper is organised as follows: In Section II, the system model is desribed. In Section III, the new protocol is introduced. Also, the error–aware ML demodulator and the low–complexity demodulator is developed. In IV, we substantiate our claims by using Monte–Carlo simulation results. Finally, Section V concludes this paper.

978-1-4673-4919-2/13/$31.00 ©2013 IEEE

2013 IEEE 18th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)

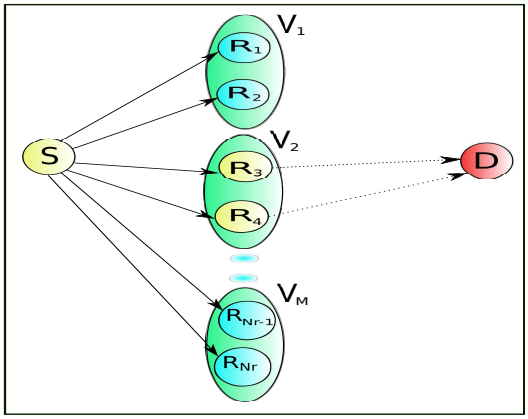


Fig. 1. DSM-STBC network topology with one source, N_r relays (partitioned into M virtual sets) and one destination. There is no coordination among any of the relays. Different line styles denote transmission over different phases: S transmits in the broadcasting phase, and the relays transmit in the relaying phase.

II. System Model and Notation

Consider a wireless network with with 1 source (S), N_r relays (R_r, with $r = 1, 2, \ldots, N_r$), and 1 destination (D), as shown in Fig. 1. Each node is equipped with a single–antenna which is subject to half duplex constraint. The channel between the source and the destination is considered to be very poor and hence, the direct link is not considered. The transmission of the data from the source to the destination takes place in two phases: 1) Broadcasting phase, where the source broadcasts its data sequentially to the N_r cooperating relays. The relays demodulate the received data using the ML criterion [25]; and 2) Relaying phase, where the relays apply the proposed Distributed Spatially–Modulated Space–Time–Block–Coding (DSM-STBC) protocol, in order to forward the received data of the source to the destination. During this phase, the source is kept silent in order to avoid receive collisions.

The fading co-efficient between node X and Y is h_{XY}, which is a circular symmetric complex Gaussian Random Variable (RV) with zero mean and variance σ^2_{XY} per dimension. The noise at the input of node Y and related to the transmission from node X is n_{XY}, which is a complex Additive White Gaussian (AWG) RV with variance $N_0/2$ per dimension. The source is assumed to transmit either Phase Shift Keying (PSK) or Quadrature Amplitude Modulation (QAM) symbols belonging to a constellation $\mathcal{A}$ with constellation size M, where $M = N_r/2$.

The N_r cooperating relays are partitioned into M virtual sets (V_s, with $s = 1, 2, \ldots, M$), where, each set consists of two distinct relays. Without loss of generality, R_1 and R_2 belong to V_1, R_3 and R_4 belong to V_2, and $R_{N_{r-1}}$ and R_{N_r} belong to V_M. We emphasize that even though the relays are partitioned into virtual sets, there is no coordination among any of the relays. Each relay is assigned a digital identifier ($\mathrm{ID}^{(R_r)}$, $r = 1, 2, \ldots, M$), of length $\log_2(M)$ bits. The relays in the same set have the same digital identifier, where as relays in different sets have different identifiers. For illustrative purposes, a lexicographic labeling is used throughout this paper. For example, if $N_r = 4$, R_1, R_2, R_3, R_4 will have identifiers, $\mathrm{ID}^{(R_1)} = \mathrm{ID}^{(R_2)} = 0$, and $\mathrm{ID}^{(R_3)} = \mathrm{ID}^{(R_4)} = 1$, respectively. In this example, the identifiers are 1 bit long since, $M = N_r/2 = 2$.

III. Proposed DSM-STBC Protocol

In this section, we introduce the proposed DSM-STBC scheme and develop an error-aware ML demodulator, which is robust to demodulation errors at the relays. We also develop a low–complexity demodulator, which simplifies the structure of the ML demodulator. Due to the low implementation constraints in applying to distributed relay networks, in this paper, we use the transmit–diversity achieving Spatially-Modulated Space-Time-Block–Codes (SM-STBC) proposed in [14]. More specifically, the so-called "set partitioning" transmission mode is used in this paper. Set partitioning mode implies that the symbols of the spatial-constellation diagram in SM are chosen such that the active antennas in each symbol result in a partition of the antenna-array. Interested readers are encouraged to refer [14] for more information.

A. Broadcasting Phase

Let s_1, s_2 and s_3 be the three data symbols that the source need to transmit to the destination, where each of s_1, s_2 and s_3 encodes $\log_2(M)$ bits. Let x_1, x_2 and x_3 be the PSK/QAM modulated symbol transmitted from the source, corresponding to s_1, s_2 and s_3, respectively. In the broadcasting phase, the source sequentially transmits x_1, x_2 and x_3 to all the cooperating relays. Thus, the signal received at the relays, R_r, for $r = 1, 2, ..., N_r$, at time–slots, $t = 1, 2, 3$ can be written as:

$$y_t^{(SR_r)} = \sqrt{E_S} h_{SR_r} x_t + n_t^{(SR_r)} \tag{1}$$

where, E_S is the source's transmit energy. The channel from the source to the r^{th} relay, h_{SR_r}, for $r = 1, 2, ..., N_r$ remains constant over the transmission of x_1, x_2 and x_3.

The signal, $y_t^{(SR_r)}$, received at each relay, R_r, for $r = 1, 2, ..., N_r$, at time–slots, $t = 1, 2, 3$ is demodulated using the ML criterion as follows:

$$\begin{cases} \hat{x}_t^{(R_r)} = \underset{\tilde{x}_t^{(R_r)} \in \mathcal{A}}{\arg\min} \left\{ \left| y_t^{(SR_r)} - \sqrt{E_S} h_{SR_r} \tilde{x}_t^{(R_r)} \right|^2 \right\} \\ \hat{s}_t^{(R_r)} = \phi_M \left(\hat{x}_t^{(R_r)} \right) \end{cases} \tag{2}$$

where: i) $\hat{x}_t^{(R_r)}$ is the source's estimated symbol at relay, R_r, at time–slot, t; ii) $\tilde{x}_t^{(R_r)}$ is the trial symbol used in the hypothesis-detection problem; iii) $\phi_M(\cdot)$ is the PSK/QAM modulation–to–bit mapping; iv) $\hat{s}_t^{(R_r)}$ is source's estimated bits.

B. Relaying Phase

After demodulating the signal received from the source using (2), each cooperating relay, R_r, for $r = 1, 2, ..., N_r$, applies the SM-STBC principle [14], to forward all the three data symbols which it received during the broadcasting phase.

The Alamouti STBC for co-located antennas is given as [1]:

$$\begin{bmatrix} x_1 & x_2 \\ -x_2^* & x_1^* \end{bmatrix} \tag{3}$$

In a co–located MIMO system with two transmit antennas, each transmit antenna transmits a distinct column of (3) in

2013 IEEE 18th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)

$$\begin{cases} y_1^{(RD)} = \sum\limits_{\Theta_r} \left(\sqrt{\frac{E_{R_r}}{2}} h_{R_r D} \hat{x}_2^{(R_r)} \Delta\left(\phi_S\left(\hat{x}_1^{(R_r)}\right), \mathrm{ID}^{(R_r)}\right) + \sqrt{\frac{E_{R_r}}{2}} h_{R_{r+1} D} \hat{x}_3^{(R_{r+1})} \Delta\left(\phi_S\left(\hat{x}_1^{(R_{r+1})}\right), \mathrm{ID}^{(R_{r+1})}\right)\right) + n_1^{(RD)} \\ y_2^{(RD)} = \sum\limits_{\Theta_r} \left(\sqrt{\frac{E_{R_r}}{2}} h_{R_r D} \left(-\hat{x}_3^{(R_r)}\right)^* \Delta\left(\phi_S\left(\hat{x}_1^{(R_r)}\right), \mathrm{ID}^{(R_r)}\right) + \sqrt{\frac{E_{R_{r+1}}}{2}} h_{R_{r+1} D} \left(\hat{x}_2^{(R_{r+1})}\right)^* \Delta\left(\phi_S\left(\hat{x}_1^{(R_{r+1})}\right), \mathrm{ID}^{(R_{r+1})}\right)\right) + n_2^{(RD)} \end{cases} \quad (4)$$

two consecutive time–slots. In DSM-STBC, Alamouti code is formed using the demodulated symbols, $\hat{x}_2^{(R_r)}$ and $\hat{x}_3^{(R_r)}$, and they are transmitted from a pair of active relays belonging a particular set, V_s, for $s = 1, 2, \ldots, M$. More specifically, $\hat{s}_1^{(R_r)}$ is used to select the relays that should be active in the relaying phase, and through these active relays, $\hat{x}_2^{(R_r)}$ and $\hat{x}_3^{(R_r)}$ are forwarded to the destination by forming the Alamouti code.

Thus, the signal received at the destination in two consecutive time–slots during the relaying phase can be written as shown in (4) on top of this page, where: i) E_{R_m} is the average total energy with which the active relay, R_m, for $m = r, r+1$, transmits in two consecutive time–slots; ii) Θ_r is defined as the set, where r takes the values $1, 3, 5, ..., N_{r-1}$; and iii) $\Delta\left(\phi_S\left(\hat{x}_1^{(R_m)}\right), \mathrm{ID}^{(R_m)}\right)$, for $m = r, r+1$, is defined as:

$$\Delta\left(\phi_S\left(\hat{x}_1^{(R_m)}\right), \mathrm{ID}^{(R_m)}\right) = \begin{cases} 1 & \text{if} \quad \hat{x}_1^{(R_m)} = \mathrm{ID}^{(R_m)} \\ 0 & \text{if} \quad \hat{x}_1^{(R_m)} \neq \mathrm{ID}^{(R_m)} \end{cases} \quad (5)$$

From (4) and (5) it follows that: i) source's data, x_1 is encoded into the activation process of the cooperating relays. More specifically, R_r, for $r = 1, 2, \ldots, N_r$, belonging to the sets V_s, for $s = 1, 2, \ldots, M$, transmits during the relaying phase if only if the demodulated data, $\hat{s}_1^{(R_m)}$, for $m = r, r+1$, obtained from (2), coincides with its own digital identifier. If this is not the case, that particular relay in silent and will have no active role to play during the relaying phase. This activation process is given by (5). In other words, x_1 is implicitly forwarded using the active relays, where as, x_2 and x_3 are forwarded using the Alamouti code from the active relays; ii) since each relay demodulates the data independently and without any coordination, some relays may decode the data correctly and some others incorrectly. Hence, the relays that are supposed to active during the relaying phase might be silent, and viceversa. The demodulation at the destination must be robust to these errors to achieve the diversity–gain provided by the Alamouti code; iii) If there are no errors at the relays, only two of the relays, R_r and R_{r+1} that belong to a particular virtual set, V_s, for $s = 1, 2, \ldots, M$, will be active during the relaying phase. In this case, conventional SM-STBC demodulator given in [14], can be used for DSM-STBC.

We can rewrite (4) alternatively as follows:

$$\mathbf{y}^{(RD)} = \sum_{\Theta_r} \mathbf{G}^{(R)} \mathbf{X}^{(R)} + \mathbf{n}^{(RD)} \quad (6)$$

where: i) $\mathbf{y}^{(RD)} = \left[y_1^{(RD)}, y_2^{(RD)}\right]$ is the received data vector at the destination; ii) $\mathbf{G}^{(R)} = \left[G^{(R_r)}, G^{(R_{r+1})}\right]$, with $G^{(R_m)} = \sqrt{\frac{E_{R_m}}{2}} h_{R_m D} \Delta\left(\phi_S\left(\hat{x}_1^{(R_m)}\right), \mathrm{ID}^{(R_m)}\right)$, for $m = r, r+1$, can be interpreted as the effective channel matrix; and iii) $\mathbf{X}^{(R)}$ is the Alamouti code transmitted by the active relays, and it is given as follows:

$$\begin{bmatrix} \hat{x}_2^{(R_r)} & \hat{x}_3^{(R_{r+1})} \\ \left(-\hat{x}_3^{(R_r)}\right)^* & \left(\hat{x}_3^{(R_{r+1})}\right)^* \end{bmatrix} \quad (7)$$

It can be seen from (7), if $\hat{x}_t^{(R_r)} \neq \hat{x}_t^{(R_{r+1})}$, for $t = 2, 3$, then, $\mathbf{X}^{(R)}\left(\mathbf{X}^{(R)}\right)^H$ is not necessarily an orthogonal STBC, which is the case when there are decoding errors at the relays. Hence, single stream decoding might not be possible for DSM-STBC, which is also the case for conventional distributed Alamouti scheme [7] .

C. Error-Aware ML Demodulator

In this section, we develop an error–aware ML demodulator for the proposed DSM-STBC scheme. This demodulator is developed by taking advantage of the knowledge of the instantaneous error probability of the source to relay links. Such kind of error-robust demodulators are necessary for achieving the diversity gain provided by the distributed Alamouti code.

By using the ML criterion [25], the 3 data symbols transmitted by the source can be jointly demodulated at the destination using (8), as shown on top of the next page, where: i) $\mathcal{P}\left\{\left[y_1^{(RD)}, y_2^{(RD)}\right] \mid \left(\tilde{\mathbf{x}}^{(D)}, \hat{\mathbf{x}}^{(D)}\right)\right\}$ is defined on top of the next page; ii) $\hat{\mathbf{x}}^{(D)} = \left[\hat{x}_1^{(D)}, \hat{x}_2^{(D)}, \hat{x}_3^{(D)}\right]$ represents the 1x3 estimated data vector of the source at the destination; iii) $\tilde{\mathbf{x}}^{(D)} = \left[\tilde{x}_1^{(D)}, \tilde{x}_2^{(D)}, \tilde{x}_3^{(D)}\right]$ represents the 1x3 trial data vector used at the destination for the hypothesis–detection problem; iv) $\hat{\mathbf{x}}^{(R)} = \left[\hat{x}_1^{(R_1)}, \hat{x}_2^{(R_1)}, \hat{x}_3^{(R_1)}, ..., \hat{x}_1^{(R_{N_r})}, \hat{x}_2^{(R_{N_r})}, \hat{x}_3^{(R_{N_r})}\right]$ represents the 1x3N_r demodulated data vector at the relays; v) $\mathcal{P}\left\{\left(\tilde{\mathbf{x}}^{(D)}, \hat{\mathbf{x}}^{(D)}\right)\right\}$ denotes the a priori joint probability of the demodulation errors at the relays, and it is defined as:

$$\mathcal{P}\left\{\left(\tilde{\mathbf{x}}^{(D)}, \hat{\mathbf{x}}^{(D)}\right)\right\} = \prod_{r=1}^{N_r} \prod_{t=1}^{3} w_t^{(R_r)}\left(\tilde{x}_t^{(D)}, \hat{x}_t^{(R_r)}\right) \quad (11)$$

where, $w_t^{(R_r)}\left(\tilde{x}_t^{(D)}, \hat{x}_t^{(R_r)}\right)$ is the probability that relay decodes the received signal to $\hat{x}_t^{(R_r)}$, when the source transmits $\tilde{x}_t^{(D)}$, at time–slot, t, of the broadcasting phase, for $t = 1, 2, 3$. $w_t^{(R_r)}\left(\tilde{x}_t^{(D)}, \hat{x}_t^{(R_r)}\right)$ is very difficult to derive for an arbitrary source signal $\tilde{x}_t^{(D)}$. Fortunately, the pair–wise error probability (PEP) of deciding $\hat{x}_t^{(R_r)}$ at the relay, when $\tilde{x}_t^{(D)}$ is transmitted from the source, denoted as $w_t^{(R_r)}\left(\tilde{x}_t^{(D)} \to \hat{x}_t^{(R_r)}\right)$, provides a tight approximation to find the solution for (8). Hence, (11) can be rewritten as follows:

$$\mathcal{P}\left\{\left(\tilde{\mathbf{x}}^{(D)}, \hat{\mathbf{x}}^{(D)}\right)\right\} = \prod_{r=1}^{N_r} \prod_{t=1}^{3} w_t^{(R_r)}\left(\tilde{x}_t^{(D)} \to \hat{x}_t^{(R_r)}\right) \quad (12)$$

Since, $w_t^{(R_r)}\left(\tilde{x}_t^{(D)} \to \hat{x}_t^{(R_r)}\right)$ is not equal to the exact error–probability $w_t^{(R_r)}\left(\tilde{x}_t^{(D)}, \hat{x}_t^{(R_r)}\right)$, the ML demodulator

2013 IEEE 18th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)

$$\begin{cases} \left[\hat{\mathbf{x}}^{(D)}\right] = \underset{\left\{\hat{x}_t^{(D)}\in\mathcal{A}\right\},t=1,2,3}{\arg\max} \left\{ \sum_{\hat{x}_1^{(R_1)}\in\mathcal{A}} \sum_{\hat{x}_2^{(R_1)}\in\mathcal{A}} \sum_{\hat{x}_3^{(R_1)}\in\mathcal{A}} \cdots \sum_{\hat{x}_1^{\left(R_{N_r}\right)}\in\mathcal{A}} \sum_{\hat{x}_2^{\left(R_{N_r}\right)}\in\mathcal{A}} \sum_{\hat{x}_3^{\left(R_{N_r}\right)}\in\mathcal{A}} \left(\mathcal{P}\left\{\left[y_1^{(RD)}, y_2^{(RD)}\right] \mid \left(\tilde{\mathbf{x}}^{(D)}, \hat{\mathbf{x}}^{(R)}\right)\right\} \mathcal{P}\left\{\left(\tilde{\mathbf{x}}^{(D)}, \hat{\mathbf{x}}^{(R)}\right)\right\}\right) \right\} \\ \hat{s}_t^{(D)} = \phi_M\left(\hat{x}_t^{(D)}\right); \quad \text{for} \quad t = 1, 2, 3 \end{cases} \tag{8}$$

$$\mathcal{P}\left\{\left[y_1^{(RD)}, y_2^{(RD)}\right] \mid \left(\tilde{\mathbf{x}}^{(D)}, \hat{\mathbf{x}}^{(R)}\right)\right\} = \exp\left(-\frac{\left|y_1^{(RD)} - \sum_{\mathcal{C}_r}\left(G^{(R_r)}\hat{x}_2^{(R_r)} + G^{\left(R_{r+1}\right)}\hat{x}_3^{\left(R_r+1\right)}\right)\right|^2}{N_0}\right) \exp\left(-\frac{\left|y_2^{(RD)} - \sum_{\mathcal{C}_r}\left(G^{(R_r)}\left(-\hat{x}_3^{(R_r)}\right)^* + G^{\left(R_{r+1}\right)}\left(\hat{x}_2^{\left(R_r+1\right)}\right)^*\right)\right|^2}{N_0}\right) \tag{9}$$

obtained by substituting (12) in (8) is not the optimal solution, but a near–optimal one. $w_t^{(R_r)}\left(\tilde{x}_t^{(D)} \to \hat{x}_t^{(R_r)}\right)$ can be derived using the union-bound approach as [25]:

$$w_t^{(R_r)}\left(\tilde{x}_t^{(D)} \to \hat{x}_t^{(R_r)}\right) = \begin{cases} Q\left(\sqrt{\frac{1}{2}\left|\tilde{x}_t^{(D)} - \hat{x}_t^{(R_r)}\right|^2 \gamma_{SR_r}}\right) & \text{if} \quad \tilde{x}_t^{(D)} \neq \hat{x}_t^{(R_r)} \\ 1 - \sum_{\substack{\hat{x}_t^{(R_r)}\in\mathcal{A} \\ \hat{x}_t^{(R_r)}\neq\tilde{x}_t^{(D)}}} Q\left(\sqrt{\frac{1}{2}\left|\tilde{x}_t^{(D)} - \hat{x}_t^{(R_r)}\right|^2 \gamma_{SR_r}}\right) & \text{if} \quad \tilde{x}_t^{(D)} = \hat{x}_t^{(R_r)} \end{cases} \tag{13}$$

where, $\gamma_{SR_r} = |h_{SR_r}|^2 (E_S/N_0)$.

Using [25, Eq. 8.26], the summation term in the second line of (13) can be avoided and a more desirable form can be obtained as shown below:

$$w_t^{(R_r)}\left(\tilde{x}_t^{(D)} \to \hat{x}_t^{(R_r)}\right) = \begin{cases} Q\left(\sqrt{\frac{1}{2}\left|\tilde{x}_t^{(D)} - \hat{x}_t^{(R_r)}\right|^2 \gamma_{SR_r}}\right) & \text{if} \quad \tilde{x}_t^{(D)} \neq \hat{x}_t^{(R_r)} \\ 1 - \beta Q\left(\sqrt{2\alpha\gamma_{SR_r}}\right) & \text{if} \quad \tilde{x}_t^{(D)} = \hat{x}_t^{(R_r)} \end{cases} \tag{14}$$

where, α and β depends on the signal constellation. For BPSK modulation, $\alpha = \beta = 1$.

The complexity of error-aware ML demodulator becomes prohibitively complex with increase in the modulation order and number of relays. In the next section, we develop a low-complexity error-aware ML demodulator by simplifying the structure of the ML demodulator given by (8), (9), and (12).

D. Low–Complexity Error Aware Demodulator

In [24], a low–complexity error aware demodulator is developed for distributed STBC using the Max-Log approximation method. In this paper, we use a different and more convenient approach to derive the the low–complexity demodulator.

First of all, we rewrite the the ML decision metric of (8) as:

$$\left[\hat{\mathbf{x}}^{(D)}, \hat{\mathbf{x}}^{(D)}\right] = \underset{\left\{\tilde{x}_t^{(D)}\in\mathcal{A}\right\},t=1,2,3}{\arg\max} \left\{ \begin{matrix} \mathcal{P}\left\{\left[y_1^{(RD)}, y_2^{(RD)}\right] \mid \left(\tilde{\mathbf{x}}^{(D)}, \hat{\mathbf{x}}^{(R)}\right)\right\} \\ \mathcal{P}\left\{\left(\tilde{\mathbf{x}}^{(D)}, \hat{\mathbf{x}}^{(R)}\right)\right\} \end{matrix} \right\} \tag{15}$$

We emphasize that the new ML decision metric of (15) is not an approximation of (8), but an exact alternative form. At large SNR values, *i.e.*, when $N_0 \to 0$, using the properties: i) $1 - Q(x) \to 1$, when $x >> 1$; and ii) The Chernoff-Bound, *i.e.*, $Q(\sqrt{x}) = \frac{1}{2}\exp\left(-\frac{x}{2}\right)$, the ML demodulator in (8) can be approximated as:

$$\left[\hat{\mathbf{x}}^{(D)}, \hat{\mathbf{x}}^{(D)}\right] = \underset{\left\{\tilde{x}_t^{(D)}\in\mathcal{A}\right\},t=1,2,3}{\arg\min} \left\{\Lambda^{(\text{High}-\text{SNR})}\left(y_1^{(RD)}, y_2^{(RD)}, \tilde{\mathbf{x}}^{(D)}, \hat{\mathbf{x}}^{(R)}\right)\right\} \tag{16}$$

where, $\Lambda^{(\text{High}-\text{SNR})}\left(y_1^{(RD)}, y_2^{(RD)}, \tilde{\mathbf{x}}^{(D)}, \hat{\mathbf{x}}^{(R)}\right)$ is defined on top of the next page. A detailed derivation is avoided due to space constraints. From (16) and (17), it can be seen that the general structure of the demodulator is now a two–dimensional array and hence searching the code set becomes less complex. However, the search space for the ML demodulator is not reduced. For large signal constellation size and number of relays, (16) and (17) still requires large number of computations. Hence, an advanced low–complexity demodulator is required to make DSM-STBC practical for real–world scenarios, which is postponed for future research.

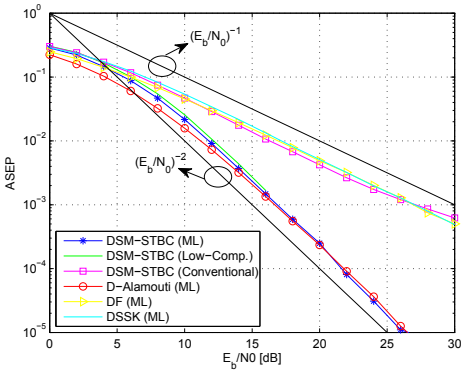


Fig. 2. Performance comparison (Monte Carlo simulations) of DSM-STBC with state-of-the-art relaying protocols. The diversity asymptotes $(E_b/N_0)^{-d}$ with diversity order d are shown as well.

IV. Simulation Results

In this section, we provide the Monte–Carlo simulation results to substantiate our claims about the performance of the proposed DSM-STBC scheme, and to compare its performance with other state–of–the–art relaying protocols. For illustrative purposes, the following setup is considered: i) $\sigma_{XY}^2 = \sigma^2 = 1/2$ for every wireless link; and ii) $E_S = E_b \log_2(M)$, $E_{R_r} = E_b\left(2\log_2(M) + \log_2\left(\lfloor N_r/2 \rfloor\right)/2\right)$, for $r = 1, 2, ...N_r$, where E_b denotes denotes the average energy per transmitted bit. For a fair comparison, all analyzed protocols have the same average transmitted energy per information bit.

In Fig. 2, the performance of the proposed DSM-STBC scheme, with $M = 2$ and $N_r = 4$ (hence, the normalized rate, $R = 3/5$ bits/CP), is compared with other state-of-the-art relaying protocols. In Fig. 2: i) "DSM-STBC (ML)" represents the DSM-STBC protocol, when the destination uses the error-aware ML demodulator in (8); ii) "DSM-STBC (Low-Comp.)"

2013 IEEE 18th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)

$$\Lambda^{(\mathrm{High-SNR})}\left(y_1^{(RD)}, y_2^{(RD)}, \hat{\mathbf{x}}^{(D)}, \hat{\mathbf{x}}^{(R)}\right) = \left|y_1^{(RD)} - \sum_{\Theta_r}\left(G^{(R_r)}\hat{x}_2^{(R_r)} + G^{\left(R_r+1\right)}\hat{x}_3^{\left(R_r+1\right)}\right)\right|^2 + \left|y_2^{(RD)} - \sum_{\Theta_r}\left(G^{(R_r)}\left(-\hat{x}_3^{(R_r)}\right)^* + G^{\left(R_r+1\right)}\left(\hat{x}_2^{\left(R_r+1\right)}\right)^*\right)\right|^2 + \sum_{r=1}^{N_r}\sum_{t=1}^{3}\left(E_S\left|h_{SR_r}\right|^2\left|\hat{x}_t^{(D)} - \hat{x}_t^{(R_r)}\right|^2\right) \quad (17)$$

is the DSM-STBC protocol, when the destination uses the low-complexity demodulator in (16); iii) "DSM-STBC (Conventional)" is the DSM-STBC protocol, when the destination uses the conventional SM-STBC protocol in [14]; iv) "D-Alamouti" is the distributed Alamouti scheme discussed in [7], with $M = 2$. An error–aware ML demodulator is used for this case as well; v) "DSSK (ML)" is the distributed space shift keying protocol proposed in [19], with $M = 2$; v) "DF (ML)" is the conventional demodulate-and-forward protocol [4], with $M = 2$.

The following comments can be made about the performance of DSM-STBC using Fig. 2: i) The DSM-STBC protocol with the error–aware ML demodulator in (8) achieves diversity order two. Moreover, the low–complexity demodulator in (16) has nearly the same performance as the ML demodulator; ii) Fig. 2 clearly shows the potential of DSM-STBC, with respect to other state–of–the–art protocols. In particular, DSM–STBC outperforms DSSK, DF and DSM-STBC (Conventional) by a huge margin, as these protocols achieve only first-order diversity gain. In terms of energy-efficiency, if $\mathrm{ASEP} = 10^{-3}$, DSM-STBC provides an RAER per transmitted bit, E_b, of approximately 90%, with respect to each of DSSK, DF and DSM-STBC (Conventional); iii) DSM-STBC has nearly the same performance as D-Alamouti, as expected. However, DSM-STBC provides higher throughput compared to D-Alamouti. More specifically, the normalized rate provided by D-Alamouti is $\frac{1}{2}\log_2(M)$ bits/CP [7], where as that of DSM-STBC is $\frac{2}{5}\log_2(M)+\frac{1}{5}\log_2(\lfloor N_r/2 \rfloor)$ bits/CP. Hence, using the set–up used in Fig. 2, the throughput provided by D-Almouti is only 1/2 bits/CP, where as that of DSM-STBC is 3/5 bits/CP. Since the proposed DSM-STBC protocol is mainly aimed at wireless networks with potentially large number of relays nodes, the throughput enhancement becomes more prominent in these scenarios. Moreover, since only 2 of these relays will be active at any given time instance, at least for high–SNR values, the total average energy consumed are the same for both protocols. Simulation results for more network set–ups are not shown here due to space constraints. The slightly better performance of D-Alamouti compared to DSM-STBC at low–SNR values is attributed to the fact that, in DSM-STBC, since there are more number of relays compared to D-Alamouti, at low–SNR, there will be more error conditions at the relays.

V. Conclusion

In this paper, DSM-STBC, a novel energy-efficient protocol for wireless networks with large number of relay nodes, is introduced. An error-aware ML and low–complexity demodulators, which are robust to demodulation errors at the relays have been developed. Monte–Carlo simulations confirmed that DSM-STBC can outperform other state–of–the–art relaying protocols.

Acknowledgment

This work is supported by the European research project "GREENET" (MITN–GA–2010–264759). The research work of M. Di Renzo is supported in part by the European Commission under the auspices of the FP7-PEOPLE MITN-CROSSFIRE project (grant 317126).

References

[1] S. M. Alamouti, "A simple transmit diversity technique for wireless communications", *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 14511458, Oct. 1998.
[2] G. J. Foschini, "Layered space–time architecture for wireless communication in fading environments when using multiple antennas", *Bell Labs*, Tech. J., vol. 2, pp. 41-59, 1996.
[3] V. Tarokh, H. Jafarkhani, and R. Calderbank, "Space-time block codes from orthogonal designs", *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 14561467, July 1999.
[4] A. Nosratinia, T. E. Hunter, and A. Hedayat, "Cooperative communications in wireless networks", *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 74–80, Oct. 2004.
[5] J. N. Laneman and G. W. Wornell, "Distributed space–time–coded protocols for exploiting cooperative diversity in wireless networks", *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2415-2425, Oct. 2003.
[6] G. Scutari, S. Barbarossa, "Distributed space–time coding for regenerative relay networks", *IEEE Trans. on Wireless Comm.*, vol. 4, no. 5, pp. 2387–2399, Sept. 2005.
[7] A. Bansal, M.R. Bhatnagar, A. Hjorungnes, "Decoding and Performance Bound of Demodulate–and–Forward Based Distributed Alamouti STBC", *IEEE Trans. on Wireless Comm.*, vol. 12, no. 2, pp. 702–713, February 2013.
[8] Simon Yiu, Robert Schober and Lutz Lampe, "Distributed space-time block coding", *IEEE Trans. on Comm.*, vol. 54, no. 07, pp. 1195–1206, July 2006.
[9] Yindi Jing and Babak Hassibi, "Distributed space–time coding in wireless relay networks", *IEEE Trans. Wireless Comm.*, vol. 5, no 12, pp. 3524–3536, Dec. 2006.
[10] R. Y. Mesleh, H. Haas, S. Sinanovic, C. W. Ahn, and S. Yun, "Spatial modulation", *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2228-2241, July 2008.
[11] M. Di Renzo, H. Haas, A. Ghrayeb, S. Sugiura, and L. Hanzo, "Spatial Modulation for Generalized MIMO: Challenges, Opportunities and Implementation", July 2013. [Online]. Available: http://hal.archives-ouvertes.fr/docs/00/84/02/78/PDF/ProcIEEE_SM_FullPaper.pdf.
[12] M. Di Renzo, H. Haas, and P. Grant, "Spatial modulation for multiple–antenna wireless systems – A survey", *IEEE Commun. Mag.*, vol. 49, pp. 182–191, Dec. 2011.
[13] M. Di Renzo and H. Haas, "Transmit-diversity for spatial modulation (SM): Towards the design of high-rate spatially-modulated space-time block codes", *IEEE Int. Commun. Conf.*, pp. 1-6, June 2011.
[14] M. Di Renzo and H. Haas, "On Transmit–Diversity for Spatial Modulation MIMO: Impact of Spatial–Constellation Diagram and Shaping Filters at the Transmitter", *IEEE Trans. on Veh. Technol.*, vol. 62, no. 6, pp. 2507–2531, July 2013.
[15] E. Basar, U. Aygolu, E. Panayirci, and H. V. Poor, "Space-time block coded spatial modulation", *IEEE Trans. Commun.*, vol. 59, no. 3, pp. 823-832, Mar. 2011.
[16] D. Yang, C. Xu, L.L. Yang, and L. Hanzo, "Transmit-diversity-assisted space-shift keying for colocated and distributed/cooperative MIMO elements", *IEEE Trans. Veh. Technol.*, vol. 60, no. 6, pp. 2864-2869, July 2011.
[17] R. Rajashekar and K. V. S. Hari, "Modulation diversity for spatial modulation using complex interleaved orthogonal design", *TENCON 2012 - 2012 IEEE Region 10 Conference*, vol. 1, no. 6, pp.19–22, Nov. 2012.
[18] S. Narayanan, M. Di Renzo, F. Graziosi, and H. Haas, "Distributed Spatial Modulation for Relay Networks", *IEEE VTC–Fall, 2013*, accepted.
[19] S. Narayanan, M. Di Renzo, F. Graziosi, and H. Haas, "Distributed Space Shift Keying for the Uplink of Relay-Aided Cellular Networks", *IEEE Int. Workshop on Computer-Aided Modeling Analysis and Design of Communication Links and Networks*, pp. 1–6, Sep. 2012.
[20] Y. Yang and S. Aissa, "Information–guided transmission in decode–and–forward relaying systems: Spatial exploitation and throughput enhancement", *IEEE Trans. Wireless Commun.*, vol. 10, no. 7, pp. 2341–2351, July 2011.
[21] S. Sugiura, S. Chen, H. Haas, P. M. Grant, and L. Hanzo, "Coherent versus non–coherent decode–and–forward relaying aided cooperative space–time shift keying", *IEEE Trans. Commun.*, vol. 59, no. 6, pp. 1707–1719, June 2011.
[22] R. Mesleh, S. Ikki, H. Aggoune, A. Mansour, "Performance Analysis of Space Shift Keying (SSK) Modulation with Multiple Cooperative Relays", *EURASIP Journal on Advances in Signal Processing*, September 2012.
[23] N. Serafimovski., S. Sinanovic., M. Di Renzo, and H. Haas, "Dual–hop spatial modulation (Dh–SM)", *IEEE Veh. Technol. Conf. – Spring*, pp. 1–5, May 2011.
[24] Chao Zhang, "Error aware distributed space–time decoding for regenerative relay networks", *IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), 2011*, vol. 35, no. 40, pp. 10–15, April 2011.
[25] M. K. Simon and M.–S. Alouini, *Digital Communication over Fading Channels*, John Wiley & Sons, 2nd ed., 2005.

F.9 IEEE Wireless Communications and Networking Conference (WCNC) 2014

IEEE WCNC'14 Track 1 (PHY and Fundamentals)

Multi-User Spatial Modulation MIMO

Sandeep Narayanan[1,2], Marium Jalal Chaudhry[1,2], Athanasios Stavridis[3],
Marco Di Renzo[4], Fabio Graziosi[2], Harald Haas[3]
[1] WEST Aquila s.r.l., 36 Strada Statale 17 Ovest, 67100 L'Aquila, Italy
[2] University of L'Aquila, Center of Excellence for Research DEWS, Via G. Gronchi 18, Nucleo Industriale di Pile, 67100 L'Aquila, Italy
[3] The University of Edinburgh, Institute for Digital Communications (IDCOM), Mayfield Road, Edinburgh, EH9 3JL, UK
[4] Laboratory of Signals and Systems (L2S), UMR 8506 CNRS – SUPELEC – Univ. Paris–Sud XI, 3 rue Joliot–Curie, 91192 Gif–sur–Yvette (Paris), France
E–Mail: {sandeep.narayanan, marium.jalal}@westaquila.com, {a.stavridis, h.haas}@ed.ac.uk, marco.direnzo@lss.supelec.fr, fabio.graziosi@univaq.it

***Abstract*— Spatial Modulation (SM) is a recently proposed single–RF multiple–input–multiple–output (MIMO) technique, which is capable of outperforming many conventional MIMO transmission schemes with low implementation and computational complexity. Recently, there have been some attempts in understanding the performance of SM in multi–user environments. However, most of the work has been oriented towards uplink multi–access scenarios. Also, conventional downlink/broadcast MIMO precoding techniques such as Zero Forcing (ZF) or Minimum Mean Square Error (MMSE) cannot be used in Multi–User SM (MU-SM), as part of the data in SM is also encoded into the Channel Impulse Responses (CIRs). In this paper, a novel precoding scheme for single–cell downlink MU-SM systems is proposed with a two–fold objective: i) the precoder needs to be able to completely eliminate the Multi–User Interference (MUI) by taking advantage of the Channel State Information (CSI) at the transmitter and ii) it needs to allow the users to use a single–user Maximum Likelihood (ML) optimum detector while achieving the same performance as interference–free point–to–point SM transmission. Finally, we also develop an interference–aware multi–user detection scheme, which does not require any CSI at the transmitter, and compare its performance with that of single–user detection schemes based on precoding.**

I. Introduction

Theoretical and practical results obtained during the past years have shown that Multiple–Input–Multiple–Output (MIMO) wireless systems can significantly increase the capacity of wireless networks [1], [2]. However, in order to satisfy the future mobile data traffic, MIMO communications will have to evolve from a single–user point–to–point transmission technique to a multi–user transmission scheme where several users can be served simultaneously using the same time and frequency resources. But, sharing the same resource block by multiple users introduces Multi–User Interference (MUI) and may reduce the true potential of MIMO systems.

Recently, several strategies have been proposed to manage the MUI in Multi–User MIMO (MU-MIMO) systems [2]–[7]. As for the uplink Multiple–Access Channel (MAC), multiple–antenna receivers at the Base Station (BS) can separate the signals coming from several different users using advanced Multi–User Detection (MUD) techniques at the cost of an increased receiver complexity [2]. As for the downlink or Broadcast Channel (BC), these complex receiver structures impose considerable design challenges at the mobile users. Previous works on MU-MIMO downlink have suggested that Dirty Paper Coding (DPC) and precoding techniques are effective solutions to eliminate or minimize the MUI [3], [4], [7]. Even though DPC can achieve the full sum capacity, it is impractical to be implemented in real networks due to its high computational complexity. On the other hand, precoding is a simple and straightforward approach for interference cancelation, provided that Channel State Information (CSI) is available at the transmitter. Some widely used precoding strategies include linear approaches based, *e.g.*, on Minimum Mean Square Error (MMSE) and Zero–Forcing (ZF) methods and non–linear approaches such as Tomlinson–Harashima precoding and vector perturbation [4]–[7] methods.

Spatial Modulation (SM) is a recently proposed single–RF MIMO technique which is capable of outperforming many conventional MIMO transmission schemes with low implementation and computational complexity [8], [9]. In SM, the data to be transmitted is encoded into two information carrying units: i) a signal constellation diagram, *i.e.*, a Phase Shift Keying (PSK)/Quadrature Amplitude Modulation (QAM) symbol and ii) a spatial constellation diagram, *i.e.*, the spatial position of a single active transmit antenna in the antenna–array. The extra degrees of freedom offered by SM based on the spatial constellation diagram introduce a multiplexing gain compared to conventional SIMO systems. Moreover, unlike spatial multiplexing MIMO systems, this multiplexing gain is obtained by activating only a single RF chain at the transmitter and by using a single–stream optimum Maximum Likelihood (ML) demodulator at the receiver. Space Shift Keying (SSK) modulation is a low–complexity form of SM, where the data is encoded only into the spatial constellation diagram [10]. The data rate offered by SSK modulation is lower than SM, but it requires less signal processing complexity.

The performance of SM is single–user and point–to–point scenarios has been studied extensively in the literature for both co–located [11]–[17] and distributed [18]–[25] setups. Recently, some experimental activities have been conducted as well [26], [27]. A comprehensive state–of–the–art survey of SM is available in [28], to which the interested reader is referred for further information. Recently, there have been some attempts in understanding the performance of SM in multi–user environments [29]–[33]. In [29], the authors study the performance of SSK modulation in the presence of multiple–access interference. More specifically, two kind of detectors, which are referred to as interference–unaware single–

978-1-4799-3083-8/14/$31.00 ©2014IEEE

user detector and interference–aware multi–user detector, are developed and analytically studied. Furthermore, the performance of SSK modulation is compared against conventional PSK/QAM–based single–antenna communications, and it is shown that SSK is capable of outperforming conventional schemes for various MIMO setups and channel conditions. In [30], the framework proposed for MU-SSK in [29] is generalized to SM. In [31], the author proposes an antenna–hopping spatial-division multiple–access scheme, which, in addition to supporting multiple–access capabilities, is capable of providing transmit–diversity gains. Using Space–Time Shift Keying (STSK), which is an extension of SM to the time domain, the authors of [32], propose a novel Orthogonal Frequency-Division Multiple–Access (OFDMA) based single-carrier multi–user scheme for frequency selective channels. The proposed scheme provides improved performance in dispersive channels while supporting multiple users in a multiple–antenna–aided wireless system. In [33], an Orthogonal Frequency-Division Multiplexing (OFDM) based downlink Multi–User SM (MU-SM) system is considered, where the data of each user is transmitted on different OFDM sub–channels. At the transmitter, linear precoding schemes such as ZF, MMSE and Block Diagonalization (BD) are applied for eliminating the MUI. Also, it is shown that OFDM-based MU-SM is capable of outperforming Alamouti-coded OFDM and V-BLAST OFDM systems in some scenarios. However, an Inverse Fast Fourier Transform (IFFT) needs to be computed at the transmitter [33, Fig. 1], which may cause high Peak-to-Average-Power-Ratio (PAPR) issues and may make the design of a power-efficient power amplifiers a challenging task.

By carefully looking at the research works conducted on MU-SM to date, it is apparent that most of the works have been oriented towards uplink multiple–access scenarios. However, to the best of our knowledge, there has been no research activities on downlink/broadcast MU–SM systems, which are capable of serving multiple users on the same time/frequency resources and of decoding them by using interference–unaware single–user detectors. More challengingly, conventional broadcast MU-MIMO precoding techniques such as ZF or MMSE cannot be used in MU-SM systems. This is because, unlike conventional MIMO systems, in SM part of the data is encoded also into the CIRs. Hence eliminating the channel matrix results in the loss of the transmitted data.

Motivated by these considerations, in the present paper a novel precoding scheme for single–cell downlink MU-SM systems is proposed with a two–fold objective: i) the precoder needs to be able to completely eliminate the MUI by taking advantage of the CSI at the transmitter and ii) it needs to allow the users to employ single–user ML-optimum detectors, while achieving the same performance as interference–free point–to–point SM transmission. Finally, we develop an interference–aware multi–user detection technique, which does not require any CSI at the transmitter and compare its performance with that of the single–user detection scheme based on precoding.

The remainder of the present paper is organized as follows. In Section II, the system model is introduced. In Section III, the proposed precoder for MU-SM is illustrated. In Section IV, the multi-user detection scheme is presented. In Section V, Monte–Carlo simulation results are shown to evaluate the performance of MU-SM. Finally, Section VI concludes this paper.

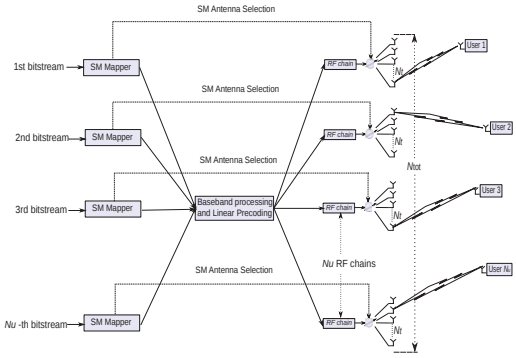


Fig. 1. Downlink MU-SM system model

II. MU-SM System Model

We consider a general single–cell downlink broadcast MU-SM setup with N_{tot} total transmit antennas and N_u active users, as shown in Fig. 1. We assume that each user is equipped with a single receive antenna. At the transmitter, the N_{tot} antennas are split into N_u blocks (b_u, for $u = 1, 2, ..., N_u$) with N_t antennas each, where $N_t = N_{tot}/N_u$. N_t is assumed to be a power of two. Without loss of generality, block b_1 is allocated to user 1, b_2 to user two, and b_{N_u} to the N_uth user. Splitting the antenna–array into blocks and allocating them to different active users is done with a two–fold objective. i) After precoding at the transmitter, each user, u for $u = 1, 2, ..., N_u$, can choose to ignore the signals coming from all the blocks other than b_u. This is particularly important in SM, as part of the information is also encoded into the CIRs of the transmit–to–receive links. ii) The decoding complexity at each user can be reduced to that of point–to–point SM transmission.

In MU–SM, random binary data, d_u, for each user u, is mapped onto the index of a single transmit antenna in block b_u as well as on a modulated symbol x_u, using an SM mapper. This kind of mapping allows us to accommodate both MU-SM and multi–user SSK (MU-SSK) during signal transmission. As for SM, x_u is taken from a PSK/QAM constellation $\mathcal{A}$ of size M and unit power constraint. As for SSK, on the other hand, x_u is equal to 1. The SM–mapped data vector is preprocessed by using an appropriate precoder, described in Section III, in order to manage the MUI.

The following assumptions and notation are used throughout this paper:

1) In all wireless links, frequency–flat independent non–identically distributed Rayleigh fading is assumed. This allows us to account for different propagation distances

and shadowing effects. We denote the CIR from the tth ($t = 1, 2, ..., N_t$) transmit antenna of the bth ($b = 1, 2, ..., N_u$) block to the uth ($u = 1, 2, ..., N_u$) user as $h_u^{(b,t)}$. Furthermore, $h_u^{(b,t)}$ is a circular symmetric complex Gaussian Random Variable (RV) with zero mean and standard deviation $\sigma_u^{(b,t)}$ per dimension.

2) Using the previously mentioned assumption that block b_1 is allocated to user 1, b_2 to user 2, and b_{N_u} to the N_uth user, the channel $h_u^{(b,t)}$ is the so–called intended link or useful link, if $u = b$, and it is an interference link, if $u \neq b$.
3) The transmitter is assumed to have perfect CSI of all the wireless links. Whereas, the receiver at each user, u for $u = 1, 2, ..., N_u$, needs only the CSI of the wireless link between itself and the N_t transmit antennas in b_u.
4) The noise n_u at the input of the receiver at user u ($u = 1, 2, ..., N_u$), is assumed to be an Additive White Gaussian Noise (AWGN) process, with both real and imaginary parts having a power spectral density equal to $N_0/2$.
5) E_m is the average energy with which each active transmit antenna transmits information. In accordance with the SM principle, only one antenna is active at any given time instance in each block. Since there are a total of N_u blocks at the transmitter, the total number of active RF chains at any time instance is N_u in MU–SM. Hence, for MU-SM, the total number of active RF chains is the same as that of conventional MU-MIMO schemes for the same number of users [4], [7].

III. Precoding Scheme for MU-SM

In this section, we propose a novel precoder design in order to eliminate the interference among the streams in an MU-SM system. The proposed precoder, which is designed by taking advantage of the CSI at the transmitter, will help the users to achieve the same performance as interference-free point-to-point SM transmission using a single-user ML-optimum detector. The precoder for MU-SM is different from the conventional MU-MIMO precoding techniques such as ZF or MMSE. This is because, in SM part of the information is also encoded in the CIRs and eliminating the channel matrix results in a loss of data.

A. Precoder Design

In MU-SM, out of the $\log_2(N_t) + \log_2(M)$ information bits for each user u ($u = 1, 2, ..., N_u$), $\log_2(N_t)$ bits are encoded into the index of a single transmit antenna in block b_u, which is switched on for data transmission, where all the other $N_t - 1$ antennas are kept silent. The remaining $\log_2(M)$ bits are encoded into a PSK/QAM symbol, which is transmitted through the antenna which is switched on. The SM–mapped data is then preprocessed using the precoder before transmission to each user. Accordingly, the signal received at the ξth user, for $\xi = 1, 2, ..., N_u$, after propagation through the wireless fading channel is given by:

$$y_\xi = \underbrace{\sqrt{E_m} p_\xi h_\xi^{(\xi,t)} x_\xi}_{\text{useful signal}} + \underbrace{\sum_{\substack{u=1 \\ u\neq\xi}}^{N_u} \left[\sqrt{E_m} p_u h_\xi^{(u,t)} x_u\right]}_{\text{interference}} + \underbrace{n_\xi}_{\text{AWGN}} \quad (1)$$

where: i) the first term in the right hand side of (1) represents the desired/useful signal received at the ξth user from the tth active antenna of the ξth antenna–block at the transmitter, ii) the second term represents the interference signal from the active antennas of all the other $N_u - 1$ blocks and iii) p_k, for $k = 1, 2, ..., N_u$, represents the precoder applied at the kth block for the interference–free reception at the kth user. For MU-SM, x_k, for $k = 1, 2, ..., N_u$, is a PSK/QAM symbol transmitted from the active antenna of the kth block. For MU-SSK, x_k, for $k = 1, 2, ..., N_u$, is equal to 1.

We can rewrite (1) in matrix form as follows:

$$\mathbf{Y} = \sqrt{E_m}\mathbf{H_{eff}}\mathbf{P} + \mathbf{n} \quad (2)$$

where: i) $\mathbf{Y} = \begin{bmatrix} y_1 & \dots & y_{N_u} \end{bmatrix}^T$ is the $\mathbb{C}^{N_u \times 1}$ received data vector; ii) $\mathbf{P} = \begin{bmatrix} p_1 & \dots & p_{N_u} \end{bmatrix}^T$ is the $\mathbb{C}^{N_u \times 1}$ precoding vector; iii) $\mathbf{n} = \begin{bmatrix} n_1 & \dots & n_{N_u} \end{bmatrix}^T$ is the $\mathbb{C}^{N_u \times 1}$ noise vector at the input of the users; and iv) $\mathbf{H_{eff}} = \begin{bmatrix} h_1^{(1,t)} x_1 & h_1^{(2,t)} x_2 & \dots & h_1^{(N_u,t)} x_{N_u} \\ h_2^{(1,t)} x_1 & h_2^{(2,t)} x_2 & \dots & h_2^{(N_u,t)} x_{N_u} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_u}^{(1,t)} x_1 & \dots & \dots & h_{N_u}^{(N_u,t)} x_{N_u} \end{bmatrix}$ is the $\mathbb{C}^{N_u \times N_u}$ effective channel–modulation matrix, where each element in the matrix is obtained by multiplying the channel gain of the active transmit antenna of each user with the QAM symbol (for MU-SM) or 1 (for MU-SSK) transmitted from that active antenna.

For interference–free reception, the required received data vector at the users needs to be equal to $\mathbf{H_{req}}$, where $\mathbf{H_{req}} = \begin{bmatrix} h_1^{(1,t)} x_1 & h_2^{(2,t)} x_2 & \dots & h_{N_u}^{(N_u,t)} x_{N_u} \end{bmatrix}^T$ is a $\mathbb{C}^{N_u \times 1}$ data vector. In formulas, the interference-free reception condition can be formulated as follows:

$$\sqrt{E_m}\mathbf{H_{eff}}\mathbf{P} = \sqrt{E_m}\mathbf{H_{req}} \quad (3)$$

From (3), the precoder can be obtained as:

$$\mathbf{P} = \mathbf{H_{eff}}^{-1}\mathbf{H_{req}} \quad (4)$$

From a closer inspection of $\mathbf{H_{eff}}$ and $\mathbf{H_{req}}$, it can be seen that $\mathbf{H_{req}}$ can be obtained from the diagonal elements of $\mathbf{H_{eff}}$. Hence, we can rewrite (4) in a more desirable form as:

$$\mathbf{P} = \mathbf{H_{eff}}^{-1}\mathrm{diag}(\mathbf{H_{eff}}) \quad (5)$$

B. Normalization of the Precoder

The normalized precoder at the transmitter can be expressed as:

$$\mathbf{G} = \bar{\beta}\mathbf{P} \quad (6)$$

where $\bar{\beta}$ is a scaling factor introduced to keep the average total power constant. The value $\bar{\beta}$ can be computed as:

$$\left[\hat{t}^{(\xi)},\hat{x}^{(\xi)}\right]=\underset{\substack{\tilde{t}\in\{1,2,\ldots,N_t\}\\ \tilde{x}_\xi\in\mathcal{A}}}{\arg\min}\left\{\sum_{\substack{\tilde{l}_u\in\{1,2,\ldots,N_t\}\\ \tilde{x}_u\in\{\mathcal{A}\}}}\left|y_\xi-\left(\sqrt{E_m}h_\xi^{(\xi,\tilde{t})}\tilde{x}_\xi+\sum_{\substack{u=1\\u\neq\xi}}^{N_u}\left[\sqrt{E_m}h_\xi^{(u,\tilde{l}_u)}\tilde{x}_u\right]\right)\right|^2\right\} \tag{12}$$

$$\bar{\beta}=\mathrm{E}\left\{\sqrt{\frac{N_u}{\mathrm{tr}\left(\mathbf{H_{eff}}^{-1}\mathrm{diag}(\mathbf{H_{eff}})\left(\mathbf{H_{eff}}^{-1}\mathrm{diag}(\mathbf{H_{eff}})\right)^{\mathrm{H}}\right)}}\right\} \tag{7}$$

where, $\mathrm{E}\{\cdot\}$ is the expectation operator.

From (7), it can be seen that the scaling factor $\bar{\beta}$ is computed by imposing an average constraint instead of an instantaneous constraint. This is because the scaling factor depends on $\mathbf{H_{eff}}$, which is a function of the channel gains from the active antennas to the users. Since some information symbols at the transmitter determine the active antennas, we cannot apply an instantaneous constraint and still provide single-stream decoding complexity. Instead, similar to the average constraint that is imposed on the data symbol (QAM/PSK) in conventional MU-MIMO systems, in MU-SM we impose an average constraint. Using this approach, $\bar{\beta}$ is just a constant that depends on the statistics of the channels, whose value can be computed at the transmitter and can then be passed on to each user.

C. Maximum–Likelihood Detector

The pre-processed data received at the users can be expressed as:

$$\begin{aligned}\mathbf{Y}&=\sqrt{E_m}\mathbf{H_{eff}}\mathbf{G}+\mathbf{n}\\&=\sqrt{E_m}\bar{\beta}\mathbf{H_{req}}+\mathbf{n}\end{aligned} \tag{8}$$

Alternatively, the received signal at each user ξ, for $\xi=1,2,...,N_u$, is as follows:

$$y_\xi=\sqrt{E_m}\bar{\beta}h_\xi^{(\xi,t)}x_\xi+n_\xi \tag{9}$$

The received signal in (9) is the same as that of the interference–free single–user SM case. Hence, an interference–unaware single–user detector [34] can be employed at the users to demodulate the data. Then, the optimal detector at each user, ξ, for $\xi=1,2,...,N_u$, based on the ML principle can be written as:

$$\left[\hat{t}^{(\xi)},\hat{x}^{(\xi)}\right]=\underset{\substack{\tilde{t}\in\{1,2,\ldots,N_t\}\\ \tilde{x}_\xi\in\mathcal{A}}}{\arg\min}\left\{\left|y_\xi-\sqrt{E_m}\bar{\beta}h_\xi^{(\xi,\tilde{t})}\tilde{x}_\xi\right|^2\right\} \tag{10}$$

where $\hat{t}^{(\xi)}$ and $\hat{x}^{(\xi)}$ are the estimated active antenna–index and the symbol from the M-ary constellation $\mathcal{A}$, respectively. As for MU-SSK, only $\hat{t}^{(\xi)}$ needs to be estimated. $\tilde{t}$ and $\tilde{x}_\xi$ are the trial instances of t and x_ξ used in the hypothesis testing problem.

IV. Multi-User detection for MU-SM

In the previous section, we developed a precoding scheme for MU-SM, which can eliminate the MUI and can help the users to achieve the same performance as interference–free SM transmission by using an interference–unaware single–user ML–optimum detector. The main advantage of the pre-coded scheme is the low computational complexity at the users, whereas its main disadvantage is that the transmitter requires CSI of all active users. In this section, we develop an interference–aware ML–optimum multi–user detector, which eliminates the need for any CSI at the transmitter.

In the absence of a precoding scheme at the transmitter, the received signal at the each user, given in (1), can be modified and written as:

$$y_\xi=\underbrace{\sqrt{E_m}h_\xi^{(\xi,t)}x_\xi}_{\text{useful signal}}+\underbrace{\sum_{\substack{u=1\\u\neq\xi}}^{N_u}\left[\sqrt{E_m}h_\xi^{(u,t)}x_u\right]}_{\text{interference}}+\underbrace{n_\xi}_{\text{AWGN}} \tag{11}$$

The signal, y_ξ, received at each user ξ, for $\xi=1,2,...,N_u$, can be demodulated by using the ML principle as shown in (12) at the top of this page, where a notation similar to Section III is used.

Some comments are worth being made about the proposed multi–user detection scheme: i) the overhead related to the acquisition of CSI can be reduced as CSI is now required only at the users, but not at the transmitter; ii) the computational complexity at the users, which is required to demodulate the transmitted data, is higher. More specifically, the complexity increases exponentially with the increase of the number of transmit antennas and the modulation order.

V. Simulation Results

In this section, we provide some simulation results to evaluate the performance of MU-SM and MU-SSK in terms of Average Symbol Error Probability (ASEP). The main purpose of the simulations is to compare the error performance of the users in the multi–user scenario with that of an interference–free single–user scenario. Both systems employ the conventional ML optimum demodulator as given in (10). The system model introduced in Section II is reproduced in our simulations. In particular: i) $\sigma_u^{(b,t)}=\sigma=1/\sqrt{2}$ is considered for all wireless links; ii) the number of transmit antennas allocated to each user, N_t, is considered to be the same for both single– and multi–user scenarios. Hence, for a single–user setup, $N_{tot}=N_t$; iii) for a fair comparison, all analyzed scenarios have the same average energy, E_m, with which the data of each user is transmitted; and iv) the scaling

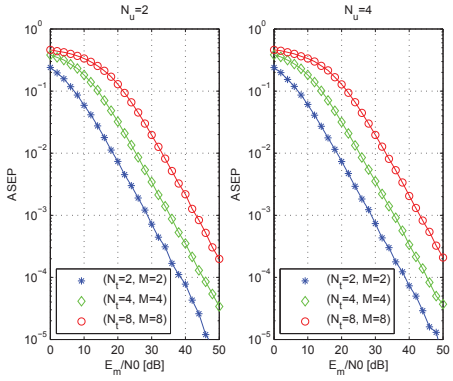


Fig. 2. ASEP of MU-SM with single–user detection [*i.e.*, (9)]. Setup: (left) $N_u = 2$ and (right) $N_u = 4$. Markers show Monte Carlo simulations for the user 1 in an MU-SM system, and solid lines show the Monte Carlo simulations for the conventional single–user SM [34].

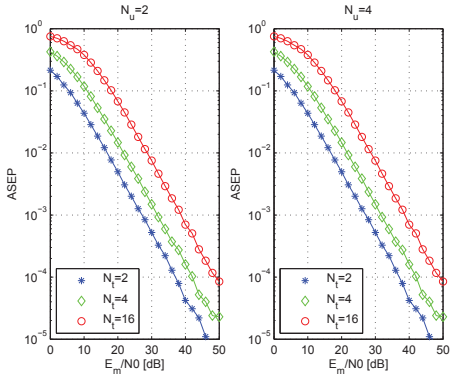


Fig. 3. ASEP of MU-SSK with single–user detection [*i.e.*, (9)]. Setup: (left) $N_u = 2$ and (right) $N_u = 4$. Markers show Monte Carlo simulations for the user 1 in an MU-SSK system, and solid lines show the Monte Carlo simulations for the conventional single–user SSK [10].

factor, $\bar{\beta}$, given by (7), is computed by generating random realizations of the i.i.d fading and signal constellation. Our extensive computer simulations show that the value of $\bar{\beta}$ is approximately equal to 1 for all the cases.

In Fig. 2, we study the robustness of the proposed precoding scheme of Section III for various MIMO setups (N_t and M) as a function of the number of active users N_u. We note that the users in a MU-SM system have the same performance as that of the single–user setup, thus confirming the ability of the precoder to totally suppress the MUI. Moreover, the ASEP of MU-SM is independent of N_u and no error-floor is present. In Fig. 3, the performance of a MU-SSK system is compared with its single–user counterpart. Our simulations show that MU-

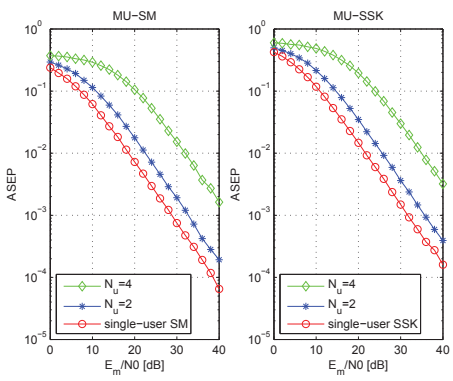


Fig. 4. ASEP of MU-SM (left) and MU-SSK (right) with multi–user detection [*i.e.*, (12)]. Setup: (left) $N_t = 2$ and $M = 2$ and (right) $N_t = 4$. Blue and green lines show Monte Carlo simulations for the user 1 in a multi–user setup with $N_u = 2$ and $N_u = 4$, respectively. Red lines show Monte Carlo simulations for the conventional single–user SM/SSK [34], [10].

SSK follows similar trends as MU-SM and thus they further confirm the potential benefits of using the proposed precoder to mitigate the interference among the streams in a multi–user scenario.

In Fig. 4, the robustness of interference-aware multi–user detection to N_u is studied and its performance is compared to the interference–free single–user detection. The setup considered is $N_t = M = 2$ for MU-SM and $N_t = 4$ for MU-SSK. It can be observed from Fig. 4 that the performance of multi–user detection without precoding is worse than that of the single–user SM/SSK transmission. Moreover, the performance of multi–user detection deteriorates further by increasing N_u. However, no error floor is present with multi–user detection and the ASEP goes to zero if the noise is very small.

By comparing Fig. 2, Fig. 3 and Fig. 4, we observe that multi–user detection is not the best choice in MU–SM/SSK. As expected, the precoded scheme with single–user detection provides better performance than multi-user detection. More specifically, by neglecting the computational complexity at the transmitter and the receiver, as well as the overhead for channel acquisition, the proposed precoding scheme allows us to reduce the average energy per transmitted bit for both MU-SM and MU-SSK. For example, if ASEP $= 10^{-3}$ the energy reduction is approximately 60% if $N_u = 2$ and 95% if $N_u = 4$, when compared to the multi–user detection scheme. For the ease of readability, the performance of only one user (user 1) is shown in all the figures. We emphasize that the performance is same for all the users.

In summary, the main aim of the present paper was to introduce SM in a multi–user environment and to propose a precoding and a multi–user detection scheme to deal with the interference. Fig. 2–Fig. 4 confirm that the proposed precoding scheme fulfills the desired goal.

F.10 IEEE 19th Int. Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)

Average Bit Error Probability of Receive-Spatial Modulation Using Zero-Forcing Precoding

Athanasios Stavridis[(1)], Dushyantha Basnayaka[(1)], Marco Di Renzo[(2)], and Harald Haas[(1)]
[(1)]Institute for Digital Communications (IDCOM), School of Engineering, The University of Edinburgh, Edinburgh, UK
[(2)]Laboratory of Signals and Systems (L2S), UMR 8506 CNRS – SUPELEC – Univ Paris–Sud, Paris, France
E–Mail: {a.stavridis, d.basnayaka, h.haas}@ed.ac.uk, marco.direnzo@lss.supelec.fr

***Abstract*—In this paper, we study the Average Bit Error Probability (ABEP) of Receive-Spatial Modulation (R-SM) in Rayleigh channels. R-SM is a multi-antenna wireless scheme that extends the concept of Spatial Modulation (SM) at the receiver side by using linear precoding. An accurate statistical framework is used for the characterization of the received signal. Using this framework, we provide novel and analytical bounds that characterize the ABEP of R-SM. Monte Carlo simulation results are used for the verification of the derived bounds. We compare R-SM against the corresponding standard Multiple-Input Multiple-Output (MIMO) scheme and show its performance. We conclude that R-SM is able to offer the same or better BER performance as traditional MIMO only when the spectral efficiency is low and the number of transmit and receive antennas is low.**

I. Introduction

The information-theoretical analysis of Multiple-Input Multiple-Output (MIMO) wireless systems demonstrates their high data rate potential [1]. However, extensive research of MIMO wireless communication systems has produced several architectures which have different characteristics. In general, MIMO systems are categorized in four groups. The first group includes systems which achieve high data rate by using parallel data streams. This is referred as spatial multiplexing. The next group subsumes schemes that are oriented to achieve diversity benefits and lower Bit Error Rate (BER) performance. Array processing gains and higher receive Signal-to-Noise Ratio (SNR) are the objectives of the third category of MIMO. An example of this category is a linearly precoded system using Channel State Information at the Transmitter (CSIT). The last group includes systems which combine characteristics from the previous groups.

Usually in MIMO, multiple Radio Frequency (RF) chains are required in order to support multiple antennas. Unfortunately, RF chains are expensive circuits that do not follow Moore's law in progressive improvement [2]. For this reason, the real system implementation becomes expensive and impractical. Moreover, multiple RF chains may reduce the Energy Efficiency (EE) of a wireless system [3]. Hence, it is important that the proposed MIMO architectures overcome the need for multiple RF chains.

A MIMO scheme that requires a single RF chain is Spatial Modulation (SM) [4, 5]. The main advantages of SM are: i) it avoids Inter-Antenna Synchronization (IAS) at the transmitter and Inter-Channel Interference (ICI) at the receiver; ii) it achieves a multiplexing gain even though it employs a single stream Maximum Likelihood (ML) detector [6]; and iii) it is able to achieve significant EE gains [3].

Based on the concept of SM and the benefits of linear precoding at the transmitter, [7] proposes a new closed-loop system, which in this paper is termed Receive-Spatial Modulation (R-SM). Using linear precoding and the appropriate transmit signal, the transmitter is able to beamform to a single receive antenna of the multi-antenna receiver. Following the SM principle, a number of bits are conveyed by the index of the receiving antenna, in addition to a number of bits encoded in a symbol drawn from a conventional constellation diagram. The BER performance of R-SM under realistic CSIT is explored in [8]. Another form of space modulation using MIMO precoding is presented in [9].

The theoretical performance analysis of R-SM is provided in [10]. In this paper, we employ the framework proposed in [10] in order to provide new analytical bounds on the Average Bit Error Probability (ABEP) of R-SM when: i) the precoding method is Zero-Forcing (ZF); ii) CSIT is perfect; and iii) the fading channel is Rayleigh. Usually, the performance analysis of ZF systems is based on a simplified statistical model which assumes uncorrelated Random Variables (RVs) in order to gain tractability [11–13]. The statistical framework of [10] avoids the previous simplification. The main contribution of this paper is the derivation of tighter bounds than [10]. The new bounds are verified by means of Monte Carlo simulations. Finally, for the purpose of comparison, we compare the BER performance of R-SM against the BER performance of the corresponding traditional MIMO scheme which: i) deploys linear precoding and ii) forms a number of parallel data steams which is equal to the number of receive antennas. We show that R-SM achieves the same BER performance as traditional MIMO only when the spectral efficiency is low and the number of transmit and receive antennas is small.

The remainder of this paper is organized as follows: Section II provides the system model of R-SM. The derivation of the analytical bounds of the ABEP of R-SM is given in Section III. Furthermore, Section IV presents the performance of R-SM and the tightness of the new bounds. Finally, conclusions are given in Section V.

Notation: In the following, lowercase bold letters denote vectors and uppercase bold letters denote matrices. Notations $(\cdot)^T$, $(\cdot)^H$, $\mathrm{tr}(\cdot)$ and $\mathbf{A}^{1/2}$ denote transpose, Hermitian transpose, matrix trace and the square root of $\mathbf{A}$, respectively. The symbol $\otimes$ stands for the Kronecker product. The Euclidean norm is denoted as $\|\cdot\|_2$. We represent a diagonal matrix, whose main diagonal includes the elements $a_1, \ldots, a_n$, as

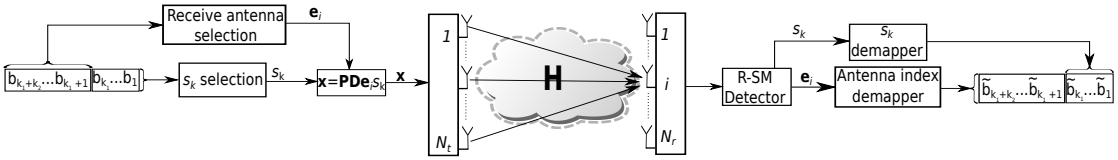


Fig. 1. The system model of Receive-Spatial Modulation.

$\mathbf{diag}\,(a_1,\ldots,a_n)$. The mean value of a RV is denoted as $\mathrm{E}[\cdot]$. A complex Gaussian distribution with mean m and variance σ_C^2 is represented as $\mathcal{CN}(m,\sigma_C^2)$, where its real and imaginary parts are independent and identically distributed (i.i.d.) Gaussian RV with distribution $\mathcal{N}(m,\frac{\sigma_C^2}{2})$. The real part of a complex scalar or matrix is denoted by $\mathrm{Re}\{\cdot\}$.

II. System Model

We consider a point-to-point MIMO communication system with N_t antennas at the transmitter and N_r antennas at the receiver. In addition, we assume Perfect-Channel State Information at the Transmitter (P-CSIT) obtained by using either the channel reciprocity or a fast and error free feedback channel. In a practical scenario, CSIT is subjected to imperfections. However, in this paper we assume P-CSIT in order to enable mathematical tractability to the presented ABEP bounds. Due to the existence of P-CSIT, the transmitter is able to employ a linear precoder. In this case, the general MIMO baseband system equation is expressed as

$$\mathbf{y} = \mathbf{HPDx} + \mathbf{w}, \tag{1}$$

where $\mathbf{y}$ is a $N_r \times 1$ vector which denotes the received signal. The channel matrix is represented as $\mathbf{H} \in \mathcal{C}^{N_r \times N_t}$. We focus on a scenario where the channel follows a Rayleigh distribution $\mathbf{H} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma}_{\mathrm{CM}})$. Here, $\mathbf{\Sigma}_{\mathrm{CM}}$ represents the covariance matrix of $\mathbf{H}$. Furthermore, $\mathbf{P}$ is a $N_t \times N_r$ matrix which denotes the linear precoder. In order to constrain the transmission power to acceptable levels, we employ the diagonal normalization matrix $\mathbf{D} = \mathbf{diag}\,(d_1,\ldots,d_{N_r})$. Each element d_i, $i = 1,\ldots,N_r$, can be found as $d_i = \sqrt{\frac{1}{\|\mathbf{p}_i\|_2^2}}$, where $\mathbf{p}_i$ represents the i-th column of $\mathbf{P}$. In this way, each column of the normalized precoding matrix $\mathbf{P}_{\mathrm{norm}} = \mathbf{PD}$ has unity power. Moreover, the transmitted signal vector is denoted as $\mathbf{x}$ and is assumed to be normalized as $\mathrm{E}_{\mathbf{x}}[\mathbf{x}] = 1$. Finally, the white Gaussian noise is represented as $\mathbf{w} \sim \mathcal{CN}(\mathbf{0}, \sigma_{\mathbf{w}}^2\mathbf{I})$, where $\sigma_{\mathbf{w}}^2$ is the variance of each element of $\mathbf{w}$.

The precoding method of interest in this paper is ZF. Hence, the linear precoder of (1), is expressed as:

$$\mathbf{P} = \mathbf{H}^H \left(\mathbf{HH}^H\right)^{-1}. \tag{2}$$

In addition, when the precoding method is ZF, the elements of the main diagonal of $\mathbf{D}$ take the following form:

$$d_i = \sqrt{\frac{1}{\left[\left(\mathbf{HH}^H\right)^{-1}\right]_{i,i}}},\ i = 1,\ldots,N_r, \tag{3}$$

where, $[\mathbf{A}]_{i,i}$ is the (i,j) element of matrix $\mathbf{A}$. Thus, if we plug (2) and (3) in (1), the received signal is formed as:

$$\mathbf{y} = \mathbf{Dx} + \mathbf{w}. \tag{4}$$

In a R-SM communication system, such as the example given in Fig. 1, during a symbol period the transmitted bitstream is encoded: i) in the index of a single receiving antenna (out of N_r); and ii) in the transmission of a constellation point drawn from a conventional constellation diagram such as M-ary Quadrature Amplitude Modulation (QAM) or Binary Phase Shift Keying (BPSK). For this reason, the R-SM transmission mode dictates that the transmitted vector $\mathbf{x}$ is written as $\mathbf{x} = \mathbf{e}_i s_k$. Here, $\mathbf{e}_i$, $i = 1,\ldots,N_r$, is a vector which is equal to the i-th column of $\mathbf{I}_{N_r,N_r}$. Hence, all the elements of $\mathbf{e}_i$ are zero except the one in the i-th position which is equal to one. Furthermore, $s_k \in \mathcal{M} = \{s_1,\ldots,M\}$ is a constellation point which belongs to a M-ary constellation diagram.

Therefore, if we set the number of receive antennas to be a power of two, the spectral efficiency of R-SM is equal to $k_{\mathrm{t}} = \log_2(M) + \log_2(N_r)$ bits per channel use (bpcu). Where, $k_1 = \log_2(M)$ bits are encoded by means of transmission of the constellation point s_k and $k_2 = \log_2(N_r)$ bits are conveyed by the selection of one (out of N_r) receiving antenna.

At the receiver side, the transmitted bitstream is reconstructed by detecting the transmitted constellation point and the index of the receiving antenna. The optimum ML detector is formed as:

$$(\tilde{\mathbf{x}}) = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{Dx}\|_2^2. \tag{5}$$

In (5), $\tilde{\mathbf{x}}$ represents the detected symbol vector. Provided that $\tilde{\mathbf{x}} = \mathbf{e}_{\tilde{i}}\tilde{s}_k$, the detection of the receiving antenna $\tilde{i}$ and the transmitted constellation point $\tilde{s}_k$ is conducted jointly.

III. Theoretical Evaluation of the Average Bit Error Probability

In this Section, we provide an upper bound of the ABEP of R-SM using the well known union bound technique [14]. Based on the union bound technique, the ABEP is upper bounded as:

$$P_{\mathrm{bit}}(\gamma) \leq \frac{1}{MN_r k_t} \sum_{\mathbf{x}} \sum_{\mathbf{x}\neq\hat{\mathbf{x}}} d(\mathbf{x}\to\hat{\mathbf{x}}) P_{\mathrm{e}}(\mathbf{x}\to\hat{\mathbf{x}},\gamma), \tag{6}$$

where $P_{\mathrm{bit}}(\gamma)$ is the ABEP for a given transmit SNR $\gamma = \frac{1}{\sigma_{\mathbf{w}}^2}$ (recall that $\mathrm{E}_{\mathbf{x}}[\mathbf{x}] = 1$). Furthermore, $P_{\mathrm{e}}(\mathbf{x}\to\hat{\mathbf{x}},\gamma)$ denotes the Pairwise Error Probability (PEP) of transmitting the symbol vector $\mathbf{x}$ and detecting the symbol vector $\hat{\mathbf{x}}$. Finally, $d(\mathbf{x}\to\hat{\mathbf{x}})$ is the number of different bits between the bit-word represented by $\mathbf{x}$ and the bit-word represented by $\hat{\mathbf{x}}$.

The computation of (6) requires the evaluation of the probability $P_{\mathrm{e}}(\mathbf{x} \rightarrow \hat{\mathbf{x}}, \gamma)$. In order to derive the previous PEP, we have to obtain the instantaneous PEP (which is conditioned on the instantaneous channel) and then average over all channel realizations.

A. Instantaneous Pairwise Error Probability

Provided that the detection process is performed using (5), a symbol error occurs when $\mathcal{E} = \left\{ \|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2^2 > \|\mathbf{y} - \mathbf{D}\hat{\mathbf{x}}\|_2^2 \right\}$. Using a straightforward elaboration on $\mathcal{E}$, we can show that $\mathcal{E} = \left\{ -\mathrm{Re}\{\mathbf{c}^H \mathbf{D} \mathbf{w}\} > \frac{\mathbf{c}^H \mathbf{D}^2 \mathbf{c}}{2} \right\}$, where $\mathbf{c} = \mathbf{x} - \hat{\mathbf{x}}$. Given the statistical distribution of $\mathbf{w}$, we can show that $-\mathrm{Re}\{\mathbf{c}^H \mathbf{D} \mathbf{w}\} \sim \mathcal{N}(0, \sigma_{\mathbf{w}}^2 \frac{\mathbf{c}^H \mathbf{D}^2 \mathbf{c}}{2})$. Hence, by taking into account that $\gamma = \frac{1}{\sigma_{\mathbf{w}}^2}$, the instantaneous PEP is expressed as:

$$P_{\mathrm{e}}(\mathbf{x} \rightarrow \hat{\mathbf{x}}, \gamma | \mathbf{D}^2) = Q\left(\sqrt{\frac{\mathbf{c}^H \mathbf{D}^2 \mathbf{c}}{2} \gamma} \right) = Q\left(\sqrt{\frac{z}{2} \gamma} \right), \quad (7)$$

where $z = \mathbf{c}^H \mathbf{D}^2 \mathbf{c}$ and $Q(\cdot)$ is the Q-function. In order to simplify (7), we have to characterize the RV z. Thus, we have to consider the different types of symbol errors in R-SM.

Like SM, R-SM faces three types of symbol error sources: i) signal, ii) space, and iii) joint errors. A signal error $\mathcal{E}_1 = \{\mathbf{x} \rightarrow \hat{\mathbf{x}} | \{s_k \rightarrow \hat{s}_k, \mathbf{e}_i \rightarrow \mathbf{e}_i\}\}$ occurs when the index of the receiving antenna is correctly detected and the conventional constellation point s_k is incorrectly detected. In contrast, the incorrect detection of the index of the receiving antenna with the concurrent correct detection of the conventional constellation point s_k results in a space error $\mathcal{E}_2 = \{\mathbf{x} \rightarrow \hat{\mathbf{x}} | \{s_k \rightarrow s_k, \mathbf{e}_i \rightarrow \mathbf{e}_{\hat{i}}\}\}$. Finally, a joint error happens when both the index of the receiving antenna and the constellation point are incorrectly detected $\mathcal{E}_3 = \{\mathbf{x} \rightarrow \hat{\mathbf{x}} | \{s_k \rightarrow \hat{s}_k, \mathbf{e}_i \rightarrow \mathbf{e}_{\hat{i}}\}\}$.

Based on the different types of symbol errors and using the structure of $\mathbf{x}$ and $\hat{\mathbf{x}}$, the RV z is written as:

$$z = \begin{cases} |s_k - \hat{s}_k|^2 d_i^2, & \text{for } \mathcal{E}_1, \quad (8a) \\ |s_k|^2 (d_i^2 + d_{\hat{i}}^2), & \text{for } \mathcal{E}_2, \quad (8b) \\ |s_k|^2 d_i^2 + |\hat{s}_k|^2 d_{\hat{i}}^2, & \text{for } \mathcal{E}_3. \quad (8c) \end{cases}$$

B. Average Pairwise Error Probability

The inspection of (7) and (8a)-(8c) shows that the derivation of the average PEP requires the marginal probability density function (pdf) of each d_i^2, $i = 1, \ldots, N_r$, for $\mathcal{E}_1$. In addition, the same equations reveal that the derivations of the PEP of $\mathcal{E}_2$ and $\mathcal{E}_3$ can be performed using the joint pdf of each pair between d_i^2 and $d_{\hat{i}}^2$.

The pdf of d_i^2 is explicitly derived in [15] as a gamma distribution with $d_i^2 \sim \mathrm{Gamma}(L, 1)$ and $L = N_t - N_r + 1$. Here, $\mathrm{Gamma}(k, \theta)$ stands for the gamma distribution with shape k and scale θ. Thus, the marginal pdf of each d_i^2 is given as:

$$f_{d_i^2}(x) = \frac{1}{\Gamma(L)} x^{L-1} e^{-x} H_0(x), \quad (9)$$

where $H_0(x)$ is the Heaviside step function, for which we have: $H_0(x) = 0$ for $x < 0$ and $H_0(x) = 1$ for $x \geq 0$. Furthermore, $\Gamma(\cdot)$ denotes the gamma function defined in [16, p. 892].

In literature, usually the RVs d_i^2 are assumed to be statistically independent in order to simplify the whole analysis [11–13]. Because this assumption is in contradiction with the structure of $d_i^2 = 1/[(\mathbf{H}\mathbf{H}^H)^{-1}]_{i,i}$, in this paper we follow a different approach. We take into account that d_i^2, $i = 1, \ldots, N_r$, are correlated gamma RVs (with marginal pdf given in (9)). Thus, as given in [17, p. 337], the joint pdf of d_i^2 and $d_{\hat{i}}^2$ is the Kibble's bivariate gamma distribution:

$$h_{d_i^2, d_{\hat{i}}^2}(x, y) = \frac{(1 - \rho_c)^{-L}}{\Gamma(L)} \sum_{k=0}^{+\infty} b_k f_k(x) p_k(y). \quad (10)$$

In (10), we have that: $b_k = \rho_c^k / \left((1 - \rho_c)^{2k} \Gamma(L + k) k! \right)$; $f_k(x) = \left[x^{L+k-1} e^{-x/(1-\rho_c)} \right] H_0(x)$; and $p_k(y) = \left[y^{L+k-1} e^{-y/(1-\rho_c)} \right] H_0(y)$. Moreover, ρ_c stands for the Pearson product-moment correlation coefficient with $\rho_c = \mathrm{E}\left[(d_i^2 - L)(d_{\hat{i}}^2 - L) \right] / L$. Given the statistical characterization of d_i^2 and $d_{\hat{i}}^2$, the estimation of ρ_c can be carried out, for certain values of N_t and N_r, by using multiple samples of the two RVs.

In order to derive the average PEP of R-SM, when a signal error $\mathcal{E}_1$ takes place, we simplify (7) using (8a) and average over all possible realizations of d_i^2 as:

$$P_{\mathcal{E}_1}(\mathbf{x} \rightarrow \hat{\mathbf{x}}) = \mathrm{E}_{d_i^2}\left[Q\left(\sqrt{\frac{|\delta|^2 d_i^2}{2} \gamma_1} \right) \right], \quad (11)$$

where $\delta = s_k - \hat{s}_k$. In this paper, an upper bound for (11) is found by considering the fact that the $Q(\cdot)$ function is tightly upper-bounded as $Q(x) \leq \frac{1}{6} e^{-2x^2} + \frac{1}{12} e^{-x^2} + \frac{1}{4} e^{\frac{-x^2}{2}}$ [18]. Hence, (11) becomes:

$$P_{\mathcal{E}_1}(\mathbf{x} \rightarrow \hat{\mathbf{x}}) \leq \mathrm{E}_{d_i^2}\left[\frac{1}{6} e^{|\delta|^2 d_i^2 \gamma} \right] + \mathrm{E}_{d_i^2}\left[\frac{1}{12} e^{\frac{|\delta|^2 d_i^2}{2} \gamma} \right] + \mathrm{E}_{d_i^2}\left[\frac{1}{4} e^{\frac{|\delta|^2 d_i^2}{4} \gamma} \right]. \quad (12)$$

The right side of (12) shows that we need to evaluate three expectations which have the form of $g_1(\alpha_1, \beta_1) = \mathrm{E}_{d_i^2}\left[\alpha_1 e^{\beta_1 d_i^2} \right]$. Here, α_1 and β_1 are positive real constants which depend on each term of the right side of (12). Thus, if we use the marginal pdf of (9), we have:

$$g_1(\alpha_1, \beta_1) = \frac{\alpha_1}{\Gamma(L)} \int_0^{+\infty} x^{L-1} e^{(\beta_1 + 1)x} \mathrm{d}x = \alpha_1 (\beta_1 + 1)^{-L}, \quad (13)$$

where in the last step of (13), we consider the integration formula from [16, p. 892, 3.351, 3]. Finally, the incorporation of (13) in (12) results in:

$$P_{\mathcal{E}_1}(\mathbf{x} \rightarrow \hat{\mathbf{x}}) \leq \frac{1}{6} \left(|\delta|^2 \gamma + 1 \right)^{-L} + \frac{1}{12} \left(\frac{|\delta|^2 \gamma}{2} + 1 \right)^{-L} + \frac{1}{2} \left(\frac{|\delta|^2 \gamma}{4} + 1 \right)^{-L}. \quad (14)$$

The average PEP of R-SM when a space error $\mathcal{E}_2$ occurs

can be bounded by following a similar procedure as before. After the simplification of (7) using (8b), we can apply the previous bound of $Q(\cdot)$ and obtain:

$$P_{\mathcal{E}_2}(\mathbf{x}\to\hat{\mathbf{x}}) \le \mathrm{E}_{d_i^2,d_i^2}\left[\frac{1}{6}e^{-|s_k|^2\left(d_i^2+d_i^2\right)\gamma}\right] + \mathrm{E}_{d_i^2,d_i^2}\left[\frac{1}{12}e^{-\frac{|s_k|^2\left(d_i^2+d_i^2\right)\gamma}{2}}\right] + \mathrm{E}_{d_i^2,d_i^2}\left[\frac{1}{4}e^{-\frac{|s_k|^2\left(d_i^2+d_i^2\right)\gamma}{4}}\right]. \quad (15)$$

By inspecting (15), it can be found that the expectations we have to deal with have the form of $g_2(\alpha_2,\beta_2) = \mathrm{E}_{d_i^2,d_i^2}\left[\alpha_2 e^{-\beta_2\left(d_i^2+d_i\right)}\right]$, where again α_2 and β_2 are positive real scalars. Given that the joint pdf of the RVs d_i^2 and d_i^2 is given in (10), after some straightforward manipulations which are omitted due to space limitation, we have:

$$g_2(\alpha_2,\beta_2) = \frac{\alpha_2\left(1-\rho_c\right)^L}{\Gamma(L)}\sum_{k=0}^{+\infty}\left[\int_0^{+\infty} e^{-\beta_2 x} f_k(x)\mathrm{d}x \times \int_0^{+\infty} e^{-\beta_2 y} p_k(y)\mathrm{d}y\right]. \quad (16)$$

If we take into account the structure of $f_k(x)$ and $p_k(y)$, both integrations in (16) can be reduced to:

$$\int_0^{+\infty} e^{-\beta_2 x} f_k(x)\mathrm{d}x = (L+k-1)!\left(\beta_2+\frac{1}{1-\rho_c}\right)^{-L-k} \quad (17)$$

$$\int_0^{+\infty} e^{-\beta_2 y} p_k(y)\mathrm{d}y = (L+k-1)!\left(\beta_2+\frac{1}{1-\rho_c}\right)^{-L-k}. \quad (18)$$

In the derivation of (17) and (18), we use again the formula from [16, p. 892, 3.351, 3] in order to attain a closed form integration. Plugging (17) and (18) into (16), and after some rearrangements and simplifications, we can show that:

$$g_2(\alpha_2,\beta_2) = \frac{\alpha_2\left(1-\rho_c\right)^{-L}}{\tau\left(\beta_2\right)^{2L}}\ {}_1F_0\left(L;\ ;\frac{\rho_c\tau\left(\beta_2\right)^{-2}}{\left(1-\rho_c\right)^2}\right), \quad (19)$$

where $\tau(\beta_2) = \beta_2 + \frac{1}{1-\rho_c}$ and ${}_1F_0(L;\ ;y) = \sum_{k=0}^{+\infty}\frac{[L]_k}{k!}y^k$ is the generalized hypergeometric function, as defined in [16, p. 1010]. Furthermore, $[L]_k = L(L+1)\ldots(L+k-1)$ is the Pochhammer symbol. Finally, we apply (19) in (15) and obtain:

$$P_{\mathcal{E}_2}(\mathbf{x}\to\hat{\mathbf{x}}) \le \frac{\left(1-\rho_c\right)^{-L}}{6\tau\left(|s_k|^2\gamma\right)^{2L}}\ {}_1F_0\left(L;\ ;\frac{\rho_c\tau\left(|s_k|^2\gamma\right)^{-2}}{\left(1-\rho_c\right)^2}\right) + \frac{\left(1-\rho_c\right)^{-L}}{12\tau\left(\frac{|s_k|^2\gamma}{2}\right)^{2L}}\ {}_1F_0\left(L;\ ;\frac{\rho_c\tau\left(\frac{|s_k|^2\gamma}{2}\right)^{-2}}{\left(1-\rho_c\right)^2}\right) + \frac{\left(1-\rho_c\right)^{-L}}{4\tau\left(\frac{|s_k|^2\gamma}{4}\right)^{2L}}\ {}_1F_0\left(L;\ ;\frac{\rho_c\tau\left(\frac{|s_k|^2\gamma}{4}\right)^{-2}}{\left(1-\rho_c\right)^2}\right). \quad (20)$$

The final step of our proof is to derive the bound of the average PEP of R-SM, when a joint symbol error $\mathcal{E}_3$ happens. Similar to the other types of symbol errors in R-SM, by combining (7) and (8c), we have:

$$P_{\mathcal{E}_3}(\mathbf{x}\to\hat{\mathbf{x}}) \le \mathrm{E}_{d_i^2,d_i^2}\left[\frac{1}{6}e^{-\left(|s_k|^2 d_i^2+|\hat{s}_k|^2 d_i^2\right)\gamma}\right] + \mathrm{E}_{d_i^2,d_i^2}\left[\frac{1}{12}e^{-\frac{\left(|s_k|^2 d_i^2+|\hat{s}_k|^2 d_i^2\right)\gamma}{2}}\right] + \mathrm{E}_{d_i^2,d_i^2}\left[\frac{1}{4}e^{-\frac{\left(|s_k|^2 d_i^2+|\hat{s}_k|^2 d_i^2\right)\gamma}{4}}\right]. \quad (21)$$

From (21), we can see that we have to deal with three terms of the form of $g_3(\alpha_3,\beta_3) = \mathrm{E}_{d_i^2,d_i^2}\left[\alpha_3 e^{-\beta_3\left(|s_k|^2 d_i^2+|\hat{s}_k|^2 d_i^2\right)\gamma}\right]$. Here, the real and positive values of α_3 and β_3 are defined from the right side of (20).

The derivation of $g_3(\alpha_3,\beta_3)$ can be obtained by following similar steps as in (16)-(19):

$$g_3(\alpha_3,\beta_3) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left[\alpha_3 e^{-\beta_3\left(|s_k|^2 d_i^2+|\hat{s}_k|^2 d_i^2\right)\gamma} \times h_{d_i^2,d_i^2}(x,y)\right]\mathrm{d}x\mathrm{d}y = \frac{\alpha_3\left(1-\rho_c\right)^{-L}}{\left(\tau_1(\beta_3)\,\tau_2(\beta_3)\right)^L} \times\ {}_1F_0\left(L;\ ;\frac{\rho_c}{\left(1-\rho_c\right)^2\tau_1(\beta_3)\,\tau_2(\beta_3)}\right), \quad (22)$$

where $\tau_1(\beta_3) = \beta_3|s_k|^2\gamma + \frac{1}{1-\rho_c}$ and $\tau_2(\beta_3) = \beta_3|\hat{s}_k|^2\gamma + \frac{1}{1-\rho_c}$. If we insert (22) into (21), we can show that the average PEP of a joint error is bounded by (23) (top of the next page).

IV. Results and Discussion

In this Section, we provide simulation results that confirm the tightness of the bound of ABEP of Section III. Furthermore, in order to have a complete insight of R-SM, we compare its BER performance against the BER performance of the corresponding traditional spatial multiplexing MIMO scheme which uses ZF precoding and forms N_r parallel symbol streams.

We consider a 2×2 ($N_t = 2$ and $N_r = 2$) system configuration which achieves a spectral efficiency of 2 bpcu and a 4×2 ($N_t = 4$ and $N_r = 2$) with 4 bpcu spectral efficiency. Furthermore, all conventional symbols, both for R-SM and traditional MIMO are drawn from a M-ary QAM or BPSK constellation diagrams. The value of M is selected in such way that guarantees the desired spectral efficiency. Finally, we assume Rayleigh fading channels.

As it can be seen from Fig. 2, the proposed upper bounds of Section III are tight and capture the nature of the behavior of the performance of R-SM in high SNR. In low SNR, there is a deviation which is a well known effect of the union bound technique [14]. Additionally, for the purpose of comparison, in the same figure, we demonstrate the performance of the corresponding traditional MIMO scheme for the same system setups. The inspection of Fig. 2 reveals that both schemes have the same performance for a 2×2 and 2 bpcu system setup.

$$P_{\mathcal{E}_3}(\mathbf{x}\to\hat{\mathbf{x}}) \leq \frac{(1-\rho_c)^{-L}}{6\left(\tau_1(1)\,\tau_2(1)\right)^L}\;{}_1F_0\left(L;\;;\frac{\rho_c}{(1-\rho_c)^2\,\tau_1(1)\,\tau_2(1)}\right)+\frac{(1-\rho_c)^{-L}}{12\left(\tau_1\left(\frac{1}{2}\right)\tau_2\left(\frac{1}{2}\right)\right)^L}$$
$$\times\;{}_1F_0\left(L;\;;\frac{\rho_c}{(1-\rho_c)^2\,\tau_1\left(\frac{1}{2}\right)\tau_2\left(\frac{1}{2}\right)}\right)+\frac{(1-\rho_c)^{-L}}{4\left(\tau_1\left(\frac{1}{4}\right)\tau_2\left(\frac{1}{4}\right)\right)^L}\;{}_1F_0\left(L;\;;\frac{\rho_c}{(1-\rho_c)^2\,\tau_1\left(\frac{1}{4}\right)\tau_2\left(\frac{1}{4}\right)}\right) \quad (23)$$

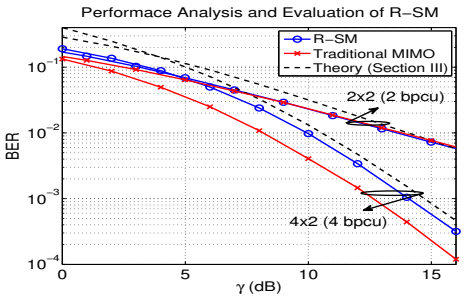


Fig. 2. Performance analysis of R-SM and its evaluation against the corresponding traditional spatial multiplexing MIMO scheme.

TABLE I
RAER GAINS OF R-SM VERSUS TRADITIONAL MIMO.

Configuration	k (bpsp)	Target BER	RAER [%]
2×2	2	10^{-2}	4.5%
4×2	4	10^{-2}	−30%

Whereas, for a system setup 4×2 and 4 bpcu traditional MIMO outperforms R-SM, due to its higher multiplexing gain.

Table I demonstrates the comparison of R-SM against traditional MIMO in terms of the Relative Average Energy Reduction (RAER) metric. RAER is defined as RAER[%] $= \left[1-10^{-\Delta_{\text{SNR}}/10}\right]\times 100\%$. Here, Δ_{SNR} denotes the SNR (in dB) difference between R-SM and traditional MIMO for a given target BER. As we see from Table I, for a 2×2 system configuration, R-SM is more energy efficient, while for a 4×2 system configuration, traditional MIMO is more energy efficient.

V. Conclusions

In this paper, we investigated the error performance of R-SM. Using an accurate statistical framework for the received signal, we derived a new upper bound for the ABEP of R-SM. We concluded that the new bound is tight. We compared R-SM against a standard spatial multiplexing MIMO system and analyzed its performance. We concluded that R-SM has the same performance as the standard MIMO only when the spectral efficiency is low and and the number of the deployed antennas is low. Thus, R-SM is not suitable for a large scale system setup. Finally, we note that the derivation of the diversity order of R-SM is a subject of our future research.

Acknowledgement

This paper has received funding from the European Union's Seventh Program for research, technological development and demonstration under grant agreement No PITN-GA-2010-264759. Prof. Harald Haas acknowledges support by EPSRC under grant EP/K008757/1.

References

[1] E. Telatar, "Capacity of Multi-Antenna Gaussian Channels," *European Trans. on Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov. / Dec. 1999.
[2] A. Molisch and M. Win, "MIMO Systems With Antenna Selection," *IEEE Microw. Mag.*, vol. 5, no. 1, pp. 46–56, Mar. 2004.
[3] A. Stavridis, S. Sinanović, M. D. Renzo., and H. Haas, "Energy Evaluation of Spatial Modulation at a Multi-Antenna Base Station," in *Proc. of the 78th IEEE Veh. Tech. Conf. (VTC)*, Las Vegas, USA, Sep. 2–5, 2013.
[4] M. Di Renzo, H. Haas, and P. M. Grant, "Spatial Modulation for Multiple-Antenna Wireless Systems: A Survey," *IEEE Commun. Mag.*, vol. 49, no. 11, pp. 182–191, Nov. 2011.
[5] M. Di Renzo, H. Haas, A. Ghrayeb, S. Sugiura, and L. Hanzo, "Spatial Modulation for Generalized MIMO: Challenges, Opportunities, and Implementation," *Proc. of the IEEE*, vol. 102, no. 1, pp. 56–103, Jan 2014.
[6] J. Jeganathan, A. Ghrayeb, and L. Szczecinski, "Spatial Modulation: Optimal Detection and Performance Analysis," *IEEE Commun. Lett.*, vol. 12, no. 8, pp. 545–547, 2008.
[7] L.-L. Yang, "Transmitter Preprocessing Aided Spatial Modulation for Multiple-Input Multiple-Output Systems," in *Proc. of 73rd IEEE Veh.Techn. Conf. (VTC Spring)*, May 2011, pp. 1 –5.
[8] A. Stavridis, S. Sinanović, M. Di Renzo, and H. Haas, "Transmit Precoding for Receive Spatial Modulation Using Imperfect Channel Knowledge," in *Proc. of 75th IEEE Veh. Techn. Conf. (VTC Spring)*, 2012.
[9] K.Ntontin, M. D. Renzo, A. I. Pérez-Neira, and C. Verikoukis, "Adaptive Generalized Space Shift Keying," *EURASIP J. on Wireless Commun. and Netw.*, 2013.
[10] A. Stavridis, D. Basnayaka, S. Sinanovic, M. Di Renzo, and H. Haas, "A Virtual MIMO Dual-Hop Architecture Based on Hybrid Spatial Modulation," *IEEE Trans. on Commun.*, vol. 62, no. 9, pp. 3161–3179, Sept 2014.
[11] X. Shao, J. Yuan, and Y. Shao, "Error Performance Analysis of Linear Zero Forcing and MMSE Precoders for MIMO Broadcast Channels," *IET Commun.*, vol. 1, no. 5, pp. 1067–1074, Oct 2007.
[12] C. Masouros, M. Sellathurai, and T. Ratnarajah, "Large-Scale MIMO Transmitters in Fixed Physical Spaces: The Effect of Transmit Correlation and Mutual Coupling," *IEEE Trans. on Commun.*, vol. 61, no. 7, pp. 2794–2804, July 2013.
[13] C.-J. Chen and L.-C. Wang, "Performance Analysis of Scheduling in Multiuser MIMO Systems with Zero-Forcing Receivers," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1435–1445, September 2007.
[14] J. G. Proakis and M. Salehi, *Communication System Engineering*. Prentice Hall, 1994.
[15] D. Gore, R. Heath, and A. Paulraj, "Transmit Selection in Spatial Multiplexing Systems," *IEEE Commun. Lett.*, vol. 6, no. 11, pp. 491–493, Nov 2002.
[16] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed., A. Jeffrey and D. Zwillinger, Eds. Academic Press, Mar. 2007, ISBN-10: 0123736374.
[17] N. Balakrishnan and C. D. Lai, *Continuous Bivariate Distributions*. Springer, 2009.
[18] M. Chiani, D. Dardari, and M. K. Simon, "New Exponential Bounds and Approximations for the Computation of Error Probability in Fading Channels," *IEEE Trans. on Wireless Commun.*, vol. 2, no. 4, pp. 840–845, July 2003.

F.11 IEEE 48th Asilomar Conf. on Signals, Systems and Computers (ASILOMAR)

Optical Spatial Modulation OFDM using Micro LEDs

Muhammad Ijaz*, Dobroslav Tsonev*, Athanasios Stavridis*, Abdelhamid Younis*
Jonathan J. D. McKendry†, Erdan Gu†, Martin D. Dawson†, Stefan Videv*and Harald Haas*

*Institute for Digital Communications
Li-Fi R&D Centre
The University of Edinburgh
Edinburgh, EH9 3JL, UK
Email: {m.ijaz, d.tsonev, a.stavridis, a.younis, h.haas}@ed.ac.uk

†Institute of Photonics, Wolfson Centre, University of Strathclyde, Glasgow, G4 0NW, U.K
Email:{jonathan.mckendry, erdan.gu, m.dawson}@strath.ac.uk

***Abstract*—This paper investigates the performance of optical spatial modulation (OSM) with orthogonal frequency division multiplexing (OFDM) in a multiple-input and multiple-output (MIMO) based visible light communication (VLC) system. Gallium nitride micro light emitting diodes (μLEDs) are used as transmitters. The performance of the OSM-OFDM system is analyzed as a function of the spatial separation, d_s and the half-power semi-angle, $\phi_{1/2}$ of the transmitter μLEDs. OSM-OFDM is compared with a standard optical spatial multiplexing (OSMX) OFDM system. The achievable data rate and the computational complexity are considered as performance metrics. The simulation results show that OSMX-OFDM can achieve higher data rates than OSM-OFDM. At the same time, however, OSM-OFDM attains a significantly lower BER and computational complexity than OSMX-OFDM.**

I. Introduction

MIMO is a well known data transmission technology which exploits the spatial dimension and has been intensively studied in radio frequency (RF) based wireless communications [1]. There is also some work on MIMO in optical wireless where it is also demonstrated that the achievable data rates can be increased [2, 3]. Spatial modulation (SM) has been introduced as a combined MIMO and digital modulation technique [4]. SM requires a low complexity detector at the receiver because inter-channel interference (ICI) is entirely avoided and no inter-antenna synchronization is required unlike in other spatial multiplexing techniques such as vertical Bell Labs layered space time (V-BLAST) MIMO [5]. SM is able to increase the spectral efficiency compared with single-input single-output (SISO) systems due to its inherent capability to transmit information in a unique way in the spatial and signaling domains [6].

In OSM, each individual light emitting diode (LED) represent a spatial constellation point that is used to carry additional information bits. The active LED is determined by the random bits at the transmitter, and it can be estimated at the receiver by using maximum-likelihood (ML) detection. This, however, requires that the channels between the transmitting and receiving LEDs are distinguishable. This becomes increasingly difficult if the separation between the LEDs is reduced – especially in an intensity modulation (IM)/direct detection (DD) system where the channel is characterised by a real-valued factor. This paper, therefore, investigates for the first time the application of SM to μLED arrays. For N_t LEDs and M-ary modulation, OSM achieves a spectral efficiency of $\log_2(MN_\mathrm{t})$ bits/s/Hz [7]. In contrast, OSMX uses each of the LEDs as an individual transmitter to convey an independent data stream. This results in parallel transmissions and consequently insignificant inter channel interference (ICI). For N_t LEDs and M-ary modulation, OSMX achieves a spectral efficiency of $N_\mathrm{t}\log_2(M)$ bits/s/Hz [8].

The performance of a VLC system depends on the light propagation characteristics of the free-space optical channel. Typically, VLC systems can be classified as diffuse or line-of-sight (LOS). In a diffuse VLC system, several paths from the optical source to the receiver exist. This makes the system robust to blockage and shadowing effects. However, path losses are typically high in this configuration and multipath propagation effects create inter symbol interference (ISI). This limits the achievable data rates [9]. In a LOS based VLC system, high data rates can be achieved, but the system is vulnerable to blockage and shadowing effects because of its directionality. Recently, a 3 Gb/s LOS VLC system based on a single 50 μm gallium nitride μLED using OFDM was successfully demonstrated [10]. The results show that for high data rates even in a LOS scenario the communication channel can be frequency selective due to the limited frequency response of LEDs.

OFDM enables low-complexity equalization and adaptive bit loading in a frequency selective channel [11]. In a first practical demonstration of OFDM applied to VLC it was demonstrated that the high peak-to-average power ratio (PAPR) in OFDM can be turned into an advantage in an IM/DD system [12]. However, clipping noise cannot be avoided completely [13]. This issue can be mitigated by the combined use of SM and OFDM as SM generates sparse OFDM frames, i.e., many subcarriers are turned off.

In this work, the bit-error ratio (BER) performance and the achievable data rates of OSM-OFDM and OSMX-OFDM are studied in an optical MIMO scenario by varying the following parameters: i) the spatial separation, d_s, between the μLEDs; and ii) the half power semi-angle, $\phi_{1/2}$ of the μLEDs transmitters. Lastly, the computational complexity measured in *bits per operation* assuming, in both cases, a ML decoder. The simulation results based on the measured non-

978-1-4799-8297-4/14/$31.00 ©2014 IEEE

flat fading channel show that data rates of several the Gb/s are possible. However, the achievable data rate is very sensitive to $d_{\rm s}$ and $\phi_{1/2}$ of the μLEDs. The results also show that the computational complexity required in OSMX-OFDM is significantly higher than in OSM-OFDM.

The rest of this paper is organized as follows: the OSM-OFDM model is introduced in Section II; the receiver complexity metric is defined in Section III; the LOS propagation model is outlined in Section IV; Section V presents the results and a discussion; and finally, conclusions are drawn in Section VI.

II. OSM-OFDM System Model

We consider a system model with $N_{\rm t}$ μLEDs at the transmitter and $N_{\rm r}$ photo-detectors at the receiver. The concept behind OSM in single-carrier transmission is to activate a single μLED during a symbol period and encode information both in the transmitted light intensity and in the index of the transmitting μLED. Here, we extend this concept to multi-carrier transmission In order to improve the utilization of the limited bandwidth of the devices and to obtain high data rates. There are several variants of optical OFDM, such as direct-current-biased optical orthogonal frequency division multiplexing (DCO-OFDM), asymmetrically clipped optical OFDM (ACO-OFDM), unipolar orthogonal OFDM (U-OFDM), and Flip-OFDM (Flip-OFDM) [14]. However, in this paper we focus on DCO-OFDM as it is demonstrated that DCO-OFDM achieves highest data rates [10].

Fig. 1 shows the block diagram of an optical wireless system which combines DCO-OFDM and SM. The system generates N_t OFDM frames of size $N{=}N_{\rm fft}/2{-}1$. Here, $N_{\rm fft}$ is the number of points of the inverse fast Fourier transform (IFFT) used in the DCO-OFDM modulator. The unique feature of the system is that in each subcarrier only one LED is active. However, each sub-carrier transfers an additional SM symbol of size $\log_2(N_t)$.

In OSM-OFDM, the transmitted bit-stream is divided into two portions. The first portion of bits, which is of length $k^i_{\rm space} = \log_2(N_{\rm t})$, is encoded via the selection of the position of the non-zero element in the subcarrier vector of size N_t, $\mathbf{s}_i$, *i.e.*, via the index of the activated μLED. The second portion of bits is encoded through the non-zero element of $\mathbf{s}_i$. The non-zero element of $\mathbf{s}_i$ is drawn as a constellation point of a M_i-ary quadrature amplitude modulation (QAM) symbol. Thus, size of the second portion of bits is $k^i_{\rm signal} = \log_2(M_i)$. Note that here the order of M_i-ary QAM can be different for the different values of i due to the non-flat VLC channel. Therefore, for a performance-optimized OFDM-based VLC system, the value of M_i is determined by an adaptive bit loading algorithm such as reported in [10]. Hence, a total of $k^i_{\rm SM} = k^i_{\rm space} + k^i_{\rm signal}$ bits are encoded by a single SM symbol vector $\mathbf{s}_i$. In this way, N SM symbol vectors $\mathbf{s}_i$, $i = 1, \cdots, N$, are obtained following the principle of SM. Each SM symbol vector $\mathbf{s}_i$ is transmitted via the i-th sub-carrier. In the next step, a matrix $\mathbf{S} = [\mathbf{s}_1, \ldots, \mathbf{s}_N]$ of size $N_{\rm t} \times N$ is constructed, by accumulating the N SM symbol vectors, in order to be transmitted by the N_t DCO-OFDM modulators. In this way, the i-th row of $\mathbf{S}$ is modulated via a DCO-OFDM modulator and transmitted from the i-th μLED.

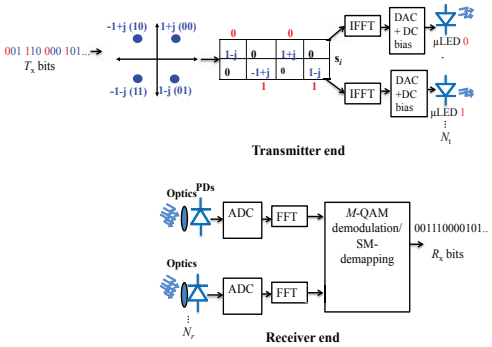


Fig. 1. Block diagram of optical SM with OFDM.

The operating principle of the DCO-OFDM modulator is as follows. In order to obtain the i-th time domain signal, the i-th row of $\mathbf{S}$ is transformed by an IFFT with $N_{\rm fft}$ points, where Hermitian symmetry is imposed in the IFFT operation [10]. After introducing a DC bias, the transformed signal is transmitted from the N_t LEDs.

At the receiver side, the transmitted optical signal is detected by the $N_{\rm r}$ photo-detectors as shown in Fig. 1. The objective of each DCO-OFDM demodulator is to transfer the signal from the time to the frequency domain with the following steps: i) the received signal is transformed by a fast Fourier transform (FFT) of $N_{\rm fft}$ points; and ii) the obsolete sub-carrier values in the resulting Hermitian-symmetric frequency-domain signal are discarded.

In this way, it is possible to accumulate the outputs of all DCO-OFDM demodulators to a matrix $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_N]$ of size $N_{\rm r} \times N$. In this study, we assume that non-linear effects in the system are insignificant. Hence, the system equation for the received SM symbol corresponding to the i-th sub-carrier can be written as follows:

$$\mathbf{y}_i = \mathbf{H}_i \mathbf{s}_i + \mathbf{w}_i. \tag{1}$$

Here, $\mathbf{y}_i$ is the i-th column of $\mathbf{Y}$. Furthermore, $\mathbf{H}_i$ denotes the optical MIMO channel gain between the $N_{\rm t}$ μLEDs and the $N_{\rm r}$ photo-detectors for the i-th sub-carrier. The effect of shot noise and thermal noise in the system is represented by $\mathbf{w}_i$. In general, $\mathbf{w}_i$ can be modelled as additive white Gaussian noise (AWGN), uncorrelated with the signal.

Provided that the system equation is given by (1), the transmitted SM symbol vectors $\mathbf{s}_i$, $i = 1, \cdots, N$, can be recovered by the following low-complexity (single stream) ML detector:

$$\hat{\mathbf{s}}_i = \arg\min_{\mathbf{s}_i} \|\mathbf{y}_i - \mathbf{H}_i \mathbf{s}_i\|_2^2, \tag{2}$$

where, $\hat{\mathbf{s}}_i$ is the detected SM symbol vector of the i-th sub-carrier. The receiver is able to reconstruct the transmitted bit-stream for the i-th sub-carrier by applying a SM de-mapper.

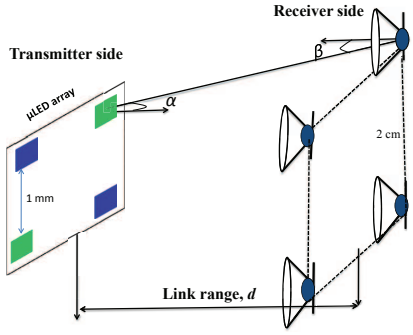


Fig. 2. Geometric scenario used to simulate the channel coefficients, **H**, in a 4×4 MIMO configuration.

III. Computational Complexity of SM ML Detection

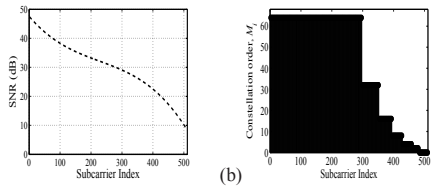
In this section, we evaluate the computational complexity of ML detection for SM. The evaluation is conducted by counting the real additions and real multiplications of the computation of the ML detection process of the i-th sub-carrier. Provided that the ML detector for SM is given in (2), if we take into account the sparsity of $\mathbf{s}_i$ (only one non-zero element) in order to reduce the number of operations, the computational complexity for the evaluation of (2) for each sub-carrier is given by:

$$\mathcal{C}^i_{\text{SM}} = 2^{k^i_{\text{SM}}} \left(16N_{\text{r}} - 2\right). \tag{3}$$

Note that the computational complexity of (2) is very low compared with the corresponding complexity of an OSMX-OFDM system with the same system setup and the same spectral efficiency. In fact, it can be shown that the computational complexity of the ML detector of an OSMX-OFDM system for the i-th sub-carrier can be given as:

$$\mathcal{C}^i_{\text{SMX}} = 2^{k^i_{\text{SMX}}} \left(8N_{\text{t}}N_{\text{r}} - 9N_{\text{r}} - 2\right), \tag{4}$$

where, k^i_{SMX} is the number of bits conveyed per sub-carrier during a signaling period. Thus, if we set $k^i_{\text{SM}} = k^i_{\text{SMX}}$, we have $\mathcal{C}^i_{\text{SM}} \ll \mathcal{C}^i_{\text{SMX}}$.



(a) (b)

Fig. 3. a) The received SNR for each sub-carrier in the non-flat VLC channel, and b) Allocated bits based on the received SNR using adaptive modulation for $\phi_{1/2} = 1^\circ$ and $d_s = 10$ mm .

IV. Communication Channel

The block diagram is presented in Fig. 2. A LOS optical wireless link is considered in the simulations of the MIMO channel, **H**. Therefore, for the given geometric scenario, the channel gain of an optical propagation link between the j-th transmitter and the i-th receiver can be calculated as follows [15]:

$$h_{(i,j)} = \begin{cases} \frac{(m+1)A}{2\pi d^2} \cos^m(\alpha) \mathrm{g}(\beta) \cos(\beta) & 0 \leq \beta \leq \psi_c \\ 0 & \beta > \psi_c \end{cases}, \tag{5}$$

where α is the angle of incidence with respect to the axis normal to the transmitter surface; β is the angle of incidence with respect to the axis normal to the receiver surface; $m = -\ln(2)/\ln(\cos(\phi_{1/2}))$ is the order of the Lambertian beam with semi-angle at half power $\phi_{1/2}$; A is the collection area of the receiver; d is the distance between the optical transmitter and the receiver; and $\mathrm{g}(\beta)$ is the concentrator gain with a half-angle field of view (FOV) of ψ_c. The main parameters used in simulating the channel gain, **H**, are given in Table I. Note that a 2×2 array of both blue and green μLEDs is used in this study as the different responsivity of the receiver to the different wavelengths can improve the distinguishability between the individual communication channels [16]. The peak wavelengths for the blue and then green μLEDs are approximately 450 nm and 510 nm, respectively. In this investigation the communication channel is not flat due to the limited frequency response of the LEDs, $\mathrm{H}(f)$, which has been measured using the setup described in [10]. Furthermore, it is assumed that the frequency responses and the optical power outputs of the green μLED and the blue μLED are same. Therefore, the received signal to noise ratio (SNR) for the communication channel between the j-th transmitter and the i-th receiver can be calculated as:

$$\mathrm{SNR}(f) = \frac{R^2 \, {\mathrm{P_t}}^2(f) \, h^2_{(i,j)} \, |\mathrm{H}(f)|^2}{w}, \tag{6}$$

where R is the detector responsivity, ${\mathrm{P_t}}^2(f)$ is the optical signal variance in a given frequency component, and w is the noise power which is calculated from the noise equivalent power (NEP) of the New Focus 1601-AC receiver used in [10].

V. Results and Discussions

In this section, the performance of an OSM-OFDM and OSMX-OFDM based VLC system is investigated in a MIMO scenario. The system block diagram for MIMO consists of 4 μLEDs and 4 photo-detectors as shown in Fig. 2. Fig. 3(a) presents an example of the received SNR, which is estimated using (6) and the experimental setup presented in [10] to

TABLE I. Main parameter used for simulating the channel gains

Parameters	Values
Half power semi-angle of μLED	1°-9°
Semi-angle, FOV of receiver	2°
Link distance	50 cm
Distance between receivers	2 cm
Area of photo-detector	0.125 mm2
Photo-detector responsivity	0.2 A/W @ 450 nm 0.3 A/W @ 550 nm
Noise equivalent power (NEP)	31 pW/$\sqrt{\text{Hz}}$
Transmit optical power	5 mW

measure the μLED frequency response, H(f). In Fig. 3(a), the received SNR values start at 47.2 dB at the low-frequency subcarriers and decrease to 8.7 dB at the high-frequency subcarriers. Thus, using a fixed QAM modulation order on all subcarriers is suboptimal. Therefore, an adaptive bit loading algorithm is employed which results in optimal QAM order allocation at the different sub-carriers according to the achievable SNR. The DCO-OFDM system uses a FFT size of N_{fft}=1024, which means that only 511 sub-carriers can be modulated with unique information due to the Hermitian symmetry requirement [10]. The single-sided bandwidth of the communication system is $B = 800$ MHz, so each sub-carrier has a bandwidth of ≈ 1.56 MHz. An example constellation size assignment is presented in Fig. 3(b) for a system configuration with $\phi_{1/2} = 1^\circ$ and $d_{\text{s}} = 10$ mm. The maximum constellation size in OSM-OFDM is set to $M = 64$, and the maximum constellation size in OSMX-OFDM is set to $M = 4$. This ensures that both systems have equivalent spectral efficiency when the OFDM subcarriers are fully loaded. Any spectral efficiency (data rate) differences between OSM-OFDM and OSMX-OFDM in the current study result from the non-flat frequency response of the μLEDs.

Fig. 4 compares the BER performance of OSM-OFDM and OSMX-OFDM as a function of the spatial separation between the μLEDs, d_{s}. The value of $\phi_{1/2}$ is set to $\phi_{1/2}$ = 1°. The spacing between the optical photo-detectors is set to 2 cm. The distance between the transmitter and the receiver receiver is 0.5 m. It can be seen that when the spatial separation between two adjacent μLEDs is less than 5 mm, both systems are unable to achieve a practical BER level of 10^{-3}. However, as the spatial separation is increased, both schemes can reach a BER of less than 10^{-3}. In fact, higher spatial separation leads to better BER performance. This behavior is explained by the fact that as the spatial separation is increased, the multiple wireless channels formed between the different μLEDs and photo-detectors are becoming more distinct. Another conclusion that can be drawn from Fig. 4 is that, for the system setup under investigation, OSM-OFDM exhibits a better BER performance than OSMX-OFDM.

Fig. 5 shows the achievable data rate as a function of the spatial separation, d_{s}, between the μLEDs. The achievable data rate is defined as $G = \sum_1^N g_i \times (1 - \text{BER}_i)$, where g_i is the error free data rate, in Gb/s, for the i-th sub-carrier, and BER_i is the achieved BER at the i-th sub-carrier. The results show that as d_{s} increases, the achievable data rate for both systems also increases. In this study, OSMX-OFDM delivers higher data rate than OSM-OFDM due to its higher multiplexing gain.

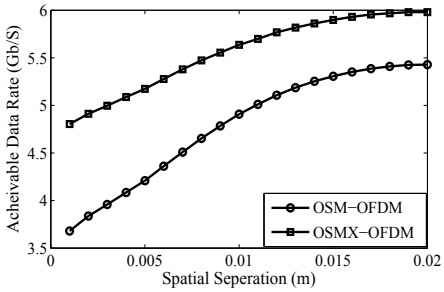


Fig. 5. Achievable data rate as a function of the transmitter separation, d_{s}.

Fig. 6 depicts the BER performance of OSM-OFDM and OSMX-OFDM as a function of $\phi_{1/2}$. In this study, d_{s}=10 mm and $\psi_{1/2}$=2°. The simulation results show that as $\phi_{1/2}$ increases, the BER performance degrades for both schemes. Clearly, the BER performance of OSM-OFDM is better than the performance of OSMX-OFDM. The corresponding achievable data rates as a function of $\phi_{1/2}$ are presented in Fig. 7. As seen from Fig. 7, OSMX-OFDM achieves higher data rates than OSM-OFDM for all values of $\phi_{1/2}$. However, for $\phi_{1/2}$>7°, OSM-OFDM and OSMX-OFDM demonstrate almost the same data rates.

The computational complexity of OSM-OFDM and OSMX-OFDM is an important measure for the practical implementation of both schemes. Fig. 8 shows the computational complexity of the ML detectors used in OSM-OFDM and OSMX-OFDM. The results in Fig. 8 have been generated according to the analysis presented in Section III. The simulation results show that the number of bits transferred per real operation at the receiver is significantly higher in OSM-

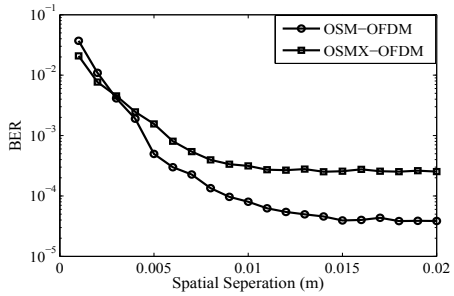


Fig. 4. BER Performance of OSM and OSMX as a function of the transmitter separation, d_{s}.

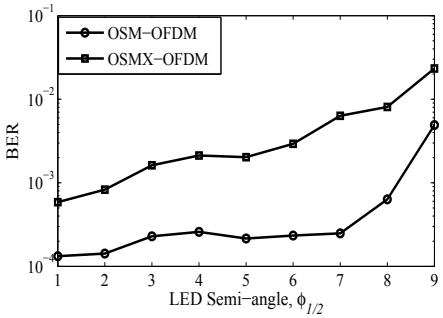


Fig. 6. BER Performance of OSM and OSMX as a function of the half power semi-angle, $\phi_{1/2}$.

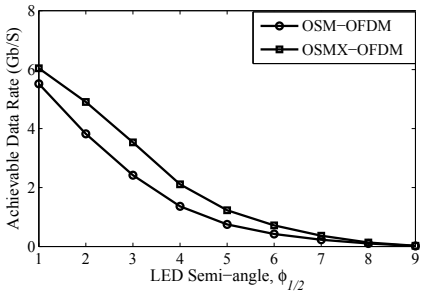


Fig. 7. Achievable data rate as a function of the transmitter half-power semi-angle, $\phi_{1/2}$.

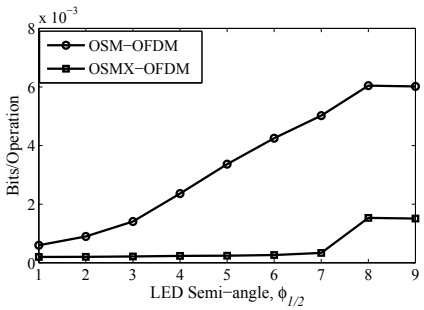


Fig. 8. Computational complexity in bits per operation as a function of the transmitter half-power semi-angle, $\phi_{1/2}$.

OFDM than in OSMX-OFDM. Therefore, when ML detection is employed, OSM-OFDM is significantly more efficient than OSMX-OFDM in terms of computational complexity.

VI. Conclusions

In this paper, we have investigated the performance of OSM-OFDM in a MIMO based VLC system. The performance of OSM-OFDM has been compared with the performance of OSMX-OFDM. The simulation results show that for the given system parameters data rates of up to 6 Gb/s are feasible using blue and green μLED devices in a spatial multiplexing MIMO configuration. The SM system achieves up to 5.5 Gb/s. This slight reduction in data rate comes with a significant reduction in the computational complexity. Furthermore, the feasibility and the performance of OSM and OSMX have been investigated as a function of the spatial separation between the μLEDs and as a function of the μLED light emission profile. The results show that in an optical MIMO scenario, the system BER performance is highly dependent on the spatial separation between the transmitters, d_s, and on the half-power semi-angle of the μLEDs, $\phi_{1/2}$. A larger spatial separation, d_s, and a smaller half power semi-angle, $\phi_{1/2}$, lead to an improved performance.

Acknowledgment

The authors gratefully acknowledge the support of this work by the UK Engineering and Physical Sciences Research Council (EPSRC) under grant EP/K00042X/1. The authors also like to thank Dr. Zheng Gong, Institute of Photonics, University of Strathclyde, Glasgow, U.K. for the fabrication of the device.

References

[1] N. Serafimovski, A. Younis, R. Mesleh, P. Chambers, M. Di Renzo, C.-X. Wang, P. Grant, M. Beach, and H. Haas, "Practical Implementation of Spatial Modulation," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 9, pp. 4511–4523, Nov 2013.

[2] L. Zeng, D. O'brien, H. Minh, G. Faulkner, K. Lee, D. Jung, Y. Oh, and E. T. Won, "High Data Rate Multiple Input Multiple Output (MIMO) Optical Wireless Communications using White Led Lighting," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 9, pp. 1654–1662, December 2009.

[3] A. Azhar, T. Tran, and D. O'Brien, "A Gigabit/s Indoor Wireless Transmission using MIMO-OFDM Visible-Light Communications," *IEEE Photonics Technology Letters*, vol. 25, no. 2, pp. 171–174, Jan 2013.

[4] R. Mesleh, H. Haas, C. W. Ahn, and S. Yun, "Spatial Modulation - A New Low Complexity Spectral Efficiency Enhancing Technique," in *International Conference on Communications and Networking in China, 2006. ChinaCom '06*, Oct 2006, pp. 1–5.

[5] P. Wolniansky, G. Foschini, G. Golden, and R. Valenzuela, "V-BLAST: An Architecture for Realizing Very High Data Rates Over the Rich-scattering Wireless Channel," in *URSI International Symposium on Signals, Systems, and Electronics, Pisa, Italy*, Sep 1998, pp. 295–300.

[6] T. Fath and H. Haas, "Performance Comparison of MIMO Techniques for Optical Wireless Communications in Indoor Environments," *IEEE Transactions on Communications*, vol. 61, no. 2, pp. 733–742, February 2013.

[7] R. Mesleh, H. Elgala, and H. Haas, "Optical Spatial Modulation," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 3, no. 3, pp. 234–244, March 2011.

[8] A. Burton, H. L. Minh, Z. Ghassemlooy, E. Bentley, and C. Botella, "Experimental Demonstration of 50 Mb/s Visible Light Communications Using 4×4 MIMO," *IEEE Photonics Technology Letters*, vol. 26, no. 9, pp. 945–948, May 2014.

[9] S. Rajbhandari, Z. Ghassemlooy, and M. Angelova, "Optimising the Performance of Digital Pulse Interval Modulation with Guard Slots for Diffuse Indoor Optical Wireless Links," *Microwaves, Antennas Propagation, IET*, vol. 5, no. 9, pp. 1025–1030, June 2011.

[10] D. Tsonev, H. Chun, S. Rajbhandari, J. McKendry, S. Videv, E. Gu, M. Haji, S. Watson, A. Kelly, G. Faulkner, M. Dawson, H. Haas, and D. O'Brien, "A 3-Gb/s Single-LED OFDM-Based Wireless VLC Link Using a Gallium Nitride μLED," *IEEE Photonics Technology Letter*, vol. 26, no. 7, pp. 637–640, April 2014.

[11] J. Armstrong, "OFDM for Optical Communications," *Journal of Lightwave Technology*, vol. 27, no. 3, pp. 189–204, Feb 2009.

[12] M. Afgani, H. Haas, H. Elgala, and D. Knipp, "Visible Light Communication using OFDM," in *2nd International Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities, 2006. TRIDENTCOM 2006*, 2006, pp. 6 pp.–134.

[13] S. Dimitrov, S. Sinanovic, and H. Haas, "Clipping Noise in OFDM-based Optical Wireless Communication Systems," *IEEE Transactions on Communications (IEEE TCOM)*, vol. 60, no. 4, pp. 1072–1081, Apr. 2012.

[14] D. Tsonev, S. Sinanovic, and H. Haas, "Complete Modeling of Nonlinear Distortion in OFDM-based Optical Wireless Communication," *Journal of Lightwave Technology*, vol. 31, no. 18, pp. 3064–3076, Sep. 15 2013.

[15] J. Kahn and J. Barry, "Wireless Infrared Communications," *Proceedings of the IEEE*, vol. 85, no. 2, pp. 265–298, Feb 1997.

[16] T. Fath and H. Haas, "Optical Spatial Modulation using Colour LEDs," in *IEEE International Conference on Communications (ICC)*, June 2013, pp. 3938–3942.

F.12 IEEE International Conference on Communications (ICC 2015)

IEEE ICC 2015 - Workshop on Visible Light Communications and Networking (VLCN)

Performance Evaluation of Space Modulation Techniques in VLC Systems

Athanasios Stavridis and Harald Haas
Li-Fi Research and Development Centre,
Institute for Digital Communications (IDCOM),
School of Engineering,
The University of Edinburgh,
Edinburgh, UK
E–Mail: {a.stavridis, h.haas}@ed.ac.uk

***Abstract*—In this paper, the Bit Error Rate (BER) performance of three major space modulation techniques in a Multiple-Input Multiple-Output (MIMO) Visible Light Communication (VLC) system is studied. The considered space modulation techniques are Optical Spatial Modulation (OSM); Optical Generalized Spatial Modulation (OGeSM); and Optical Multi-Stream Spatial Modulation (OMS-SM). The space modulation techniques are evaluated against two benchmark systems: Optical Spatial MultipleXing (OSMX) and Optical Repetition Coding (ORC). The performance assessment, for both the space modulation schemes and the benchmark systems, is undertaken using simulation and analytical results. For the considered system setup, it is concluded that, in relative low Signal-to-Noise Ratio (SNR), OSM offers the best performance. Whereas, in relative high SNR and for high spectral efficiency, OMS-SM is the most efficient scheme in terms of BER.**

I. Introduction

Visible Light Communication (VLC) is a new means of wireless communication that has the potential to provide ultra high data rates [1, 2]. Recent results demonstrate that a single Light Emitting Diode (LED) is able to provide a data rate of 3 Gb/s [3]. In addition, the incorporation of Multiple-Input Multiple-Output (MIMO) techniques in a VLC system is shown to provide further improvements in the data rate [4].

The deployment of MIMO techniques in VLC is mainly inspired by the high data rate potential of MIMO systems in Radio Frequency (RF) communication [5]. However, the techniques of RF MIMO communication are not directly transferable to VLC. In fact, the nature of the optical channel presents different challenges [6]. Hence, the research of the performance of MIMO communication in VLC is important.

In this research area, several authors have studied the deployment of MIMO techniques in VLC. In [4], the concept of Spatial MultipleXing (SMX) in VLC is presented. Furthermore, in [7], the Bit Error Rate (BER) performance of several MIMO schemes in VLC is studied. Specifically, in [7], the performance of Optical Spatial Modulation (OSM) is compared against the corresponding performance of Optical Spatial Multiplexing (OSMX) and Optical Repetition Coding (ORC). In addition, the performance of OSM is also researched in [8, 9]. The incorporation of MIMO techniques that utilize Channel State Information at the Transmitter (CSIT) in VLC is presented in [10–12].

Spatial Modulation (SM) is a successful MIMO scheme in RF communication which also has been extensively studied in VLC [7, 10–13]. Due to its operating principle, SM promotes a lower complexity transceiver implementation compared to traditional MIMO schemes, such as SMX [13, 14]. Especially, at the transmitter side only one RF chain is required. This is shown to provide energy efficiency advantages [14]. Inspired by the concept of SM, several extensions of SM have been developed. For a complete introduction of the concept of SM and its variants, the reader is referred to [13].

The objective of this paper is to extend the main space modulation schemes from RF communication to VLC. Specifically, the performance evaluation of SM [13], Generalized Spatial Modulation (GeSM) [15], and Multi-Stream-Spatial Modulation (MS-SM) [16] is studied using the metric of BER. In this paper, these schemes are termed as OSM, Optical Generalized Spatial Modulation (OGeSM), and Optical Multi-Stream-Spatial Modulation (OMS-SM), respectively. Their performance is compared against the corresponding performance of two benchmark systems: ORC and OSMX. In addition, a general theoretical framework that assesses analytically the Average Bit Error Probability (ABEP) of both the studied space modulated techniques and benchmark systems is proposed. This framework is based on the union bound technique [17]. Finally, for the considered system setup, it is concluded that OSM offers the best BER performance in relative low Signal-to-Noise Ratio (SNR). However, as the spectral efficiency and SNR are increased, OMS-SM is shown to be the most efficient scheme in terms of BER.

The remainder of this paper is organized as follows: The system model of the considered VLC system is given in Section II. In addition, Sections II-A1 to II-A3 introduce the major space modulation techniques studied in this paper. The theoretical framework that assesses the ABEP of the considered space modulation schemes and benchmark systems is presented in Section III. The analytical and simulation results that evaluate the BER performance of the different VLC transmission schemes are discussed in Section IV. Finally, the concluding remarks are given in Section V.

Notation: In the following, lowercase bold letters denote vectors and uppercase bold letters denote matrices. Notation $(\cdot)^T$ denotes the transpose of a matrix. The Euclidean norm is

978-1-4673-6305-1/15/$31.00 ©2015 IEEE

IEEE ICC 2015 - Workshop on Visible Light Communications and Networking (VLCN)

denoted as $\|\cdot\|_2$. The representation of the natural logarithm is given as $\ln(\cdot)$. A Gaussian distribution with mean m and variance σ^2 is represented as $\mathcal{N}(m, \sigma^2)$.

II. System Model

In this section, a VLC system which incorporates N_t LEDs and N_r Photo Detectors (PDs) is considered. Due to the nature of the Optical Wireless Channel (OWC), Intensity Modulation (IM) and Direct Detection (DD) are deployed. Usually, in VLC systems, only the Line-of-Sight (LOS) (dominant) component of the channel gain is considered [6, 7]. Therefore, the optical MIMO system equation is given as:

$$\mathbf{y} = r\mathbf{H}\mathbf{x} + \mathbf{w}, \tag{1}$$

where, $\mathbf{y}$ is the $N_r \times 1$ received signal vector. The responsivity of the PDs, in A/W, is denoted as r. Furthermore, $\mathbf{H}$ is a $N_r \times N_t$ matrix which denotes the optical channel. In more detail, the (i, j) element of $\mathbf{H}$, $i = 1, \ldots, N_r$ and $j = 1, \ldots, N_t$, which is denoted as $h_{i,j}$, represents the optical channel impulse response between the i-th receive PD and the j-th transmit LED. In addition, $\mathbf{x}$ is the $N_t \times 1$ transmitted signal vector. Each element of $\mathbf{x}$ is a positive number and represents the optical intensity transmitted from the corresponding LED. In order: i) to provide a fair comparison between the different transmission techniques; and ii) to ensure the efficient operation of the LEDs under the applied lighting constraints, the normalization of $\mathrm{E}_{\mathbf{x}}[\mathbf{x}] = P_{\mathrm{o}}$ is imposed. Here, P_{o} is the average optical transmission power. The way that the elements of $\mathbf{x}$ are selected depends on the deployed MIMO transmission technique. More detail is given in Section II-A, where all of the studied space modulation transmission techniques in this paper are presented. The composite effect of the ambient light shot and thermal noise is represented by $\mathbf{w}$. Following the assumptions of [6], $\mathbf{w}$ is modeled as real Additive White Gaussian Noise (AWGN), where $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma^2_{\mathbf{w}}\mathbf{I})$. Here, it holds that $\sigma^2_{\mathbf{w}} = \sigma^2_{\mathrm{shot}} + \sigma^2_{\mathrm{thermal}}$, where σ^2_{shot} and $\sigma^2_{\mathrm{thermal}}$ denote the variance of the shot and thermal noise, respectively.

As noted, this paper focuses on a LOS VLC scenario, where only the dominant component of the channel gain is considered. Therefore, according to [6], the channel impulse response between the i-th PD and the j-th LED, $h_{i,j}$, is written as:

$$h_{i,j} = \begin{cases} \frac{A(k+1)}{2\pi d^2_{i,j}} \cos^k(\phi_{i,j}) \cos(\psi_{i,j}), & 0 \le \psi_{i,j} \le \Psi_{\frac{1}{2}}, \\ 0, & \psi_{i,j} > \Psi_{\frac{1}{2}}. \end{cases} \tag{2}$$

In (2), A is the area of each PD. Furthermore, the Lambertian factor k, which determines the directionality order, is given as:

$$k = \frac{-\ln(2)}{\ln\left(\cos\left(\Phi_{\frac{1}{2}}\right)\right)}, \tag{3}$$

where, $\Phi_{\frac{1}{2}}$ denotes the transmitter semi-angle. The distance between the i-th PD and the j-th LED is represented as $d_{i,j}$. Furthermore, $\phi_{i,j}$ is the angle of emission of the j-th LED to the i-th PD with respect to the orthonormal vector of the transmitter plane of the j-th LED. In addition, $\psi_{i,j}$ represents the angle of incidence of the light at the i-th PD from the j-th LED with respect to the orthonormal vector of the receiver plane of the i-th PD. Provided that the LEDs and PDs are placed in a three dimensional Cartesian space, their positions are described by their Cartesian coordinates. The Cartesian coordinates of the j-th LED, $j = 1, \ldots, N_t$, are given from a 3×1 vector which is denoted as $\mathbf{p}^j_t$, while its orientation is given from an orthonormal vector $\mathbf{o}^j_t$ which is vertical to the plane of the LED. In the same way, the Cartesian coordinates of the i-th PD, $i = 1, \ldots, N_r$, are given from a 3×1 vector $\mathbf{p}^i_r$ and its orientation is described from an orthonormal vector $\mathbf{o}^i_r$ which is vertical to the plane of the PD. Therefore, according to [18], $\cos(\phi_{i,j})$ and $\cos(\psi_{i,j})$ can be computed as:

$$\cos(\phi_{i,j}) = \frac{\mathbf{o}^j_t\left(\mathbf{p}^i_r - \mathbf{p}^j_t\right)}{d_{i,j}} \tag{4}$$

and

$$\cos(\psi_{i,j}) = \frac{\mathbf{o}^i_r\left(\mathbf{p}^j_t - \mathbf{p}^i_r\right)}{d_{i,j}}. \tag{5}$$

Finally, the Field of View (FOV) semi-angle of every PD is denoted as $\Psi_{\frac{1}{2}}$.

At the receiver side, DD is utilized as the most practical down-conversion technique. In this case, the optimum Maximum Likelihood (ML) detector of the studied optical MIMO schemes can be expressed as:

$$(\tilde{\mathbf{x}}) = \arg\min_{\mathbf{x}} \|\mathbf{y} - r\mathbf{H}\mathbf{x}\|^2_2. \tag{6}$$

In (6), $\tilde{\mathbf{x}}$ is the detected symbol vector. Provided that $\tilde{\mathbf{x}}$ is detected at the receiver, the transmitted bit-stream can be reconstructed via the deployment of the appropriate demapping process.

A. Optical Space Modulation Techniques

This subsection introduces the operating principles and the main characteristics of the optical space modulation transmission techniques considered in this paper. The optical transmission techniques that are introduced in this subsection are: OSM; OGeSM; and OMS-SM. Note that (1) describes all the previous schemes by using the appropriate design of the transmitted vector $\mathbf{x}$. The following subsections give the design of $\mathbf{x}$ for each considered space modulation technique.

1) Optical Spatial Modulation: Similar to conventional SM in RF communication [13], the main objective of OSM is to promote low complexity system implementation at both communicating ends.

The detailed description of the operating mechanism of OSM is given below. During a symbol period, the transmitted bit-stream is divided into two sequences. The first sequence is composed from $k^{\mathrm{OSM}}_{\mathrm{Space}} = \log_2(N_t)$ bits. At this point, implicitly it is assumed that the number of LEDs is a power of two. In contrast, the length of the second sequence is $k^{\mathrm{OSM}}_{\mathrm{Signal}} = \log_2(M)$, where M is the order of the deployed IM Pulse Amplitude Modulation (PAM) constellation, $\mathcal{M}_{\mathrm{OSM}} = \{s_1, \ldots, s_M\}$. Here, s_k, $k = 1 \ldots, M$, denote the different levels of light intensity transmitted by a LED during the transmission period. Note, that due the operating principle of OSM, none of s_k, $k = 1 \ldots, M$, can have a zero value because

it corresponds to zero intensity light transmission. Otherwise, the zero value would imply the inactivity of a LED, which, as shown below, disregards the OSM transmission principle. In OSM, the first sequence of bits is encoded in the activation of one LED (out of N_r). All the other LEDs remain inactive. Provided that each LED is allocated a unique binary index of length of $k_{\text{Space}}^{\text{OSM}}$, the active LED is the one that possesses the binary index which is equal to the first sequence of bits. The second sequence of bits is encoded in the light intensity transmission of the previously selected LED. Therefore, the spectral efficiency of OSM is $k_{\text{OSM}} = k_{\text{Space}}^{\text{OSM}} + k_{\text{Signal}}^{\text{OSM}}$ bits per channel use (bpcu).

Mathematically, an OSM symbol vector is defined as:

$$\mathbf{x}_{\text{OSM}} = \mathbf{e}_i s_k, \tag{7}$$

where, $\mathbf{e}_i$ is the i-th column of the identity matrix $\mathbf{I}_{N_t,N_t} = [\mathbf{e}_1, \ldots, \mathbf{e}_{N_t}]$. The zero elements of $\mathbf{e}_i$ correspond to the inactive LEDs and the non-zero element corresponds to the active LED. In addition, s_k is the light intensity transmitted from the active LED. At the receiver side, the transmitted OSM symbol vector $\mathbf{x}_{\text{OSM}}$ is detected using (6). In this way, the transmitted bit-stream is reconstructed from the receiver.

Due to its operating principle, OSM requires one transmission chain at the transmitter. In addition, the receiver is able to deploy a low complexity (single stream) ML detector. However, despite the deployment of a single stream detector, OSM has the potential to achieve a multiplexing gain at the expense of additional LEDs.

2) Optical Generalized Spatial Modulation: As described in Section II-A, OSM requires the number of LEDs to be a power of two. However, this constraint is too restrictive. In addition, the activation of a single LED limits the number of encoded bits in only $k_{\text{Space}}^{\text{SM}}$ bits. A solution can be given via the deployment of OGeSM. Note that OGeSM is the incorporation of GeSM in VLC. The GeSM for RF communication is proposed and studied in [15].

In OGeSM, during the signaling period, instead of activating a single LED like OSM, N_a LEDs are active. Here, it holds that $1 < N_a < N_t$. In this way, binary information can be encoded in the combination of the active LEDs. Provided that N_t LEDs are available, from which only N_a are active during a symbol period, a total of $N_c = \binom{N_t}{N_a}$ combinations of active LEDs exists. Note that $\binom{\cdot}{\cdot}$ denotes the binomial coefficient. However, from the N_c combinations, only the $2^{\lfloor \log_2(N_c) \rfloor}$ can be used in order to encode binary information. The selection of the combinations which represent binary information can be done intelligently or randomly. The intelligent selection of the encoded combination can be based on a metric which minimizes the system BER. However, this method results in an additional complexity overhead. This paper, for simplicity, focuses on the random selection of the combinations of active LEDs.

Given that the combinations of active LEDs which encode binary information are selected and each combination is allocated a unique binary index, a total of $k_{\text{Space}}^{\text{OGeSM}} = \log_2\left(2^{\lfloor \log_2(N_c) \rfloor}\right)$ bits are transmitted via the index of the combination of active LEDs. In OGeSM, all of the active LEDs transmit the same light intensity which corresponds to a point drawn from a M-ary IM PAM constellation, $s_k \in \mathcal{M}_{\text{OGeSM}} = \{s_1, \ldots, s_M\}$. Similar to OSM, s_k, $k = 1, \ldots, M$, cannot take a zero value as this corresponds to zero intensity light transmission. Thus, $k_{\text{Signal}}^{\text{OGeSM}} = \log_2(M)$ bits are conveyed to the receiver through the transmission of the standard PAM point s_k. In this way, the spectral efficiency of OGeSM equals to $k_{\text{OGeSM}} = k_{\text{Space}}^{\text{OGeSM}} + k_{\text{Signal}}^{\text{OGeSM}}$ bpcu.

The mathematical formulation of an OGeSM symbol vector is given as:

$$\mathbf{x}_{\text{OGeSM}} = \mathbf{i}_{\text{OGeSM}} s_k, \tag{8}$$

where, $\mathbf{i}_{\text{OGeSM}}$ is a $N_t \times 1$ vector which represents the combination of active LEDs. Note that $\mathbf{i}_{\text{OGeSM}}$ has exactly N_a non-zero elements which are equal to one. All the other elements of $\mathbf{i}_{\text{OGeSM}}$ have a zero value. The position of a non-zero element of $\mathbf{i}_{\text{OGeSM}}$ corresponds to the position of an active LED.

The structure of an OGeSM symbol vector reveals that at the transmitter only one transmission chain is required. Indeed, the same transmission chain can drive all of the active LEDs during the signaling period (because all of the active LEDs transmit the same light intensity). Therefore, the complexity of the transmitter is not affected significantly by the use of OGeSM compared with the use of OSM. However, at the receiver side, the joint inspection of (6) and (8) shows that there is an increase in complexity compared with OSM. This happens because $\mathbf{x}_{\text{OGeSM}}$ is less sparse than $\mathbf{x}_{\text{OSM}}$.

3) Optical Multi-Stream Spatial Modulation: The spectral efficiency of OGeSM can be further increased, if each active LED transmits a different level of light intensity. In this way, a scheme is formed which spatially modulates multiple data streams from the transmitter to the receiver. This scheme is called OMS-SM and is an extension of MS-SM [16] in optical communication. The operating mechanism of OMS-SM determines that during the signaling period a combination of N_a LEDs is activated in order to encode binary information. Therefore, using the same explanation as Section II-A2, it is shown that OMS-SM encodes $k_{\text{Space}}^{\text{OMS-SM}} = \log_2\left(2^{\lfloor \log_2(N_c) \rfloor}\right)$ bits, where $N_c = \binom{N_t}{N_a}$, in the index of the combination of active LEDs. In OMS-SM, each active LED is able to transmit a different level of light intensity. Hence, every active LED is transmitting a different IM PAM symbol, $s_k \in \mathcal{M}_{\text{OMS-SM}} = \{s_1, \ldots, s_M\}$. Here, M stands for the order of the IM PAM constellation $\mathcal{M}_{\text{OMS-SM}}$. In this way, $k_{\text{Signal}}^{\text{OMS-SM}} = N_a \log_2(M)$ bits are conveyed via the N_a PAM points. Thus, the spectral efficiency of MS-SM is $k_{\text{OMS-SM}} = k_{\text{Space}}^{\text{OMS-SM}} + k_{\text{Signal}}^{\text{OMS-SM}}$ bpcu.

The mathematical description of a symbol vector $\mathbf{x}_{\text{OMS-SM}}$ of OMS-SM is given in (9) at the top of the next page. The length of $\mathbf{x}_{\text{OMS-SM}}$ is N_t elements. The i-th element of $\mathbf{x}_{\text{OMS-SM}}$ corresponds to the i-th LED. The operating principle of OMS-SM dictates that $\mathbf{x}_{\text{OMS-SM}}$ has exactly N_a non-zero elements. All the other elements equal to zero. The position of the non-zero elements correspond to the combination of active LEDs during the signaling period. The values of the non-zero elements of $\mathbf{x}_{\text{OMS-SM}}$ represent the light intensity (PAM symbols) transmitted from the corresponding

IEEE ICC 2015 - Workshop on Visible Light Communications and Networking (VLCN)

$$\mathbf{x}_{\mathrm{OMS-SM}} = \Big[0,\ldots,0,\ \underbrace{s_1}_{i_1\text{-th position}},0,\ldots,0,\ \underbrace{s_i}_{i_k\text{-th position}},0,\ldots,0,\ \underbrace{s_{N_a}}_{i_{N_a}\text{-th position}}\ 0\ldots,0\Big]^T \tag{9}$$

LEDs.

At the receiver side, during a symbol period, the transmitted bit-stream is reconstructed via the detection of the combination of active LEDs and the detection of the N_a PAM points. This is done by deploying the minimization process of (6). Note that the search of (6) is done over all possible symbol vectors of OMS-SM.

The complexity of OMS-SM is higher compared to the complexity of OSM and OGeSM. At the transmitter, N_a communication chains are required in order to produce the different levels of light intensity. Further, at the receiver side, the detection complexity is increased due to the N_a spatially modulated data streams. However, due to the sparsity of (9), it is emphasized that the complexity of an OMS-SM transceiver is lower than the corresponding complexity of a fully spatially multiplexed VLC system. In OSMX, exactly N_t parallel data streams are transmitted during the signaling period.

III. Theoretical Average Bit Error Probability

Section III provides a general theoretical framework which can be used for the evaluation of the ABEP of: OSM; OGeSM; and OMS-SM. This framework is based on the union bound technique [17]. Note that this framework can be easily extended to include the evaluation of the ABEP of any other point-to-point optical MIMO technique.

The union bound technique expresses the ABEP of a point-to-point optical MIMO communication system as:

$$P_{\mathrm{bit}}(\gamma_{\mathrm{e}}) \leq \frac{1}{|\mathcal{B}|k_t}\sum_{\mathbf{x}}\sum_{\substack{\hat{\mathbf{x}}\\ \hat{\mathbf{x}}\neq\mathbf{x}}} d(\mathbf{x}\rightarrow\hat{\mathbf{x}})P_{\mathrm{e}}(\mathbf{x}\rightarrow\hat{\mathbf{x}},\gamma_{\mathrm{e}}). \tag{10}$$

In (10), $P_{\mathrm{bit}}(\gamma_{\mathrm{e}})$ is the ABEP for a given *transmit* electrical SNR. The *transmit* electrical SNR of a VLC system is defined as $\gamma_{\mathrm{e}} = P_o^2/\sigma_{\mathbf{w}}^2$. Without loss of generality and for simplicity, here, it is assumed that the optical transmitted power P_o is normalized to unity (P_o=1). Using this form of normalization, the comparison between the different transmission techniques becomes compact as long as the same normalization is assumed. Obviously, a different normalization results in the same SNR shift for all studied transmission schemes. In addition, $\mathcal{B}$ denotes the transmission alphabet (set of all possible transmitted symbol vectors) of a certain transmission scheme, while $|\mathcal{B}|$ is the size (number of all possible transmitted symbol vectors) of the certain transmission alphabet. Furthermore, k_t denotes the number of bits transmitted per channel use. The Pairwise Error Probability (PEP) of transmitting $\mathbf{x}$ and detecting erroneously $\hat{\mathbf{x}}$, for a given value of γ_{e}, is denoted as $P_{\mathrm{e}}(\mathbf{x}\rightarrow\hat{\mathbf{x}},\gamma_{\mathrm{e}})$. Finally, $d(\mathbf{x}\rightarrow\hat{\mathbf{x}})$ is the number of different bits (Hamming distance) between the bit-word represented by $\mathbf{x}$ and the bit-word represented by $\hat{\mathbf{x}}$.

The inspection of (10) reveals that the assessment of the ABEP requires the evaluation of the PEP between all possible pairs of $\mathbf{x}$ and $\hat{\mathbf{x}}$. In the following, the derivation of the previous PEP is presented. Provided that the detection process is conducted using (6), a symbol error takes place when:

$$\mathcal{E}(\mathbf{x},\hat{\mathbf{x}}) = \left\{\|\mathbf{y}-r\mathbf{H}\mathbf{x}\|_2^2 > \|\mathbf{y}-r\mathbf{H}\hat{\mathbf{x}}\|_2^2\right\}. \tag{11}$$

After a straightforward elaboration of (11), $\mathcal{E}(\mathbf{x},\hat{\mathbf{x}})$ can be re-written as:

$$\mathcal{E}(\mathbf{x},\hat{\mathbf{x}}) = \left\{-\sum_{i=1}^{N_r}\sum_{j=1}^{N_t}\mathbf{w}_i h_{i,j}\mathbf{c}_i > \frac{r\|\mathbf{H}\mathbf{c}\|_2^2}{2}\right\}, \tag{12}$$

where, $\mathbf{c} = \mathbf{x}-\hat{\mathbf{x}}$ and $\mathbf{c}_i$, $i = 1,\ldots,N_t$, is the i-th element of $\mathbf{c}$. Given that $\mathbf{w}_i$, $i = 1,\ldots,N_r$, is the i-th element of $\mathbf{w}$ $\left(\mathbf{w}_i \sim \mathcal{N}(0,\sigma_{\mathbf{w}}^2)\right)$, it holds that:

$$-\sum_{i=1}^{N_r}\sum_{j=1}^{N_t}\mathbf{w}_i h_{i,j}\mathbf{c}_i \sim \mathcal{N}\left(0,\sigma_{\mathbf{w}}^2\|\mathbf{H}\mathbf{c}\|_2^2\right). \tag{13}$$

Therefore, using the statistical description of the previous Random Variable (RV), it is shown that the PEP of the pair of $\mathbf{x}$ and $\hat{\mathbf{x}}$ is given as:

$$P_{\mathrm{e}}(\mathbf{x}\rightarrow\hat{\mathbf{x}},\gamma_{\mathrm{e}}) = Q\left(\sqrt{\frac{\|\mathbf{H}\mathbf{c}\|_2^2}{4}r^2\gamma_{\mathrm{e}}}\right), \tag{14}$$

where, $Q(\cdot)$ is the Q-function. Provided that the Q-function is tightly upper-bounded as [19]:

$$Q(x) \leq \frac{1}{6}e^{-2x^2} + \frac{1}{12}e^{-x^2} + \frac{1}{4}e^{\frac{-x^2}{2}}, \tag{15}$$

the PEP of (14) can be expressed as in (16), at the top of the next page.

Note that in VLC systems the optical wireless channel is deterministic and does not include any randomness. In fact, multi-path fading is not present due to the size of the detector which is larger than a wavelength [6]. Thus, in contrast to RF communication, there is no need for averaging (16) over the optical channel (which has only one realization for a certain system setup).

In the final remark of Section III, it is emphasized that the ABEP of OSM, OGeSM, and OMS-SM is directly obtained from (10) via the use of (16). This is done by setting the appropriate values for $|\mathcal{B}|$ and k_t. For each considered transmission scheme, the values for the previous quantities are given in detail in Sections II-A1 to II-A3.

IV. Results and Discussion

This section provides Monte Carlo simulation results that assess the performance of: OSM; OGeSM; and OMS-SM. In addition, the simulation results are verified using the bounds of the theoretical analysis of Section III.

$$P_{\mathrm{e}}(\mathbf{x}\to\hat{\mathbf{x}},\gamma_{\mathrm{e}}) \leq \frac{1}{6}e^{-\frac{\|\mathbf{Hc}\|_2^2}{2}r^2\gamma_{\mathrm{e}}} + \frac{1}{12}e^{-\frac{\|\mathbf{Hc}\|_2^2}{4}r^2\gamma_{\mathrm{e}}} + \frac{1}{4}e^{-\frac{\|\mathbf{Hc}\|_2^2}{8}r^2\gamma_{\mathrm{e}}}. \quad (16)$$

TABLE I
COORDINATES OF LEDS AND PDS.

Transmitter Coordinates (in m)				Receiver Coordinates (in m)			
	x-axis	y-axis	z-axis		x-axis	y-axis	z-axis
LED 1	2.2	1.8	3.5	PD 1	2.15	1.85	0.85
LED 2	1.8	1.8	3.5	PD 2	1.85	1.85	0.85
LED 3	1.8	2.2	3.5	PD 3	1.85	2.15	0.85
LED 4	2.2	2.2	3.5	PD 4	2.15	2.15	0.85

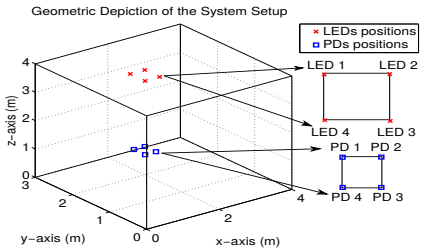
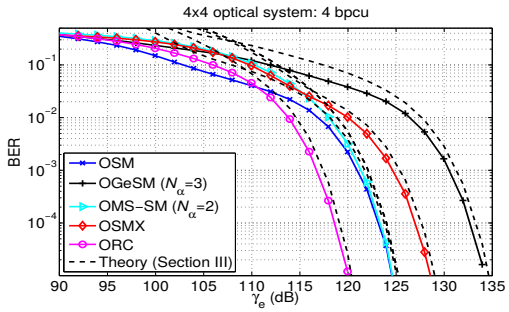


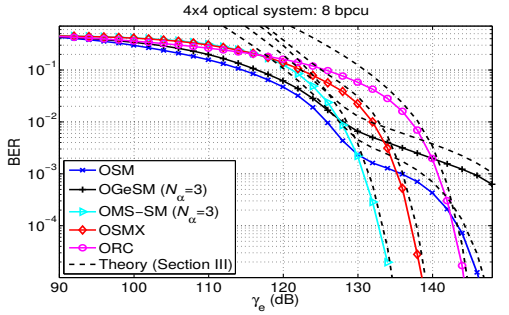
Fig. 1. Visual representation of the considered 4×4 VLC system.

For the purpose of comparison, two benchmark systems are considered. The first benchmark system is OSMX. In an OSMX system, the transmitter conveys N_t parallel data streams to the receiver. Therefore, every symbol period, binary information is transmitted via N_t points drawn from a M-ary IM PAM constellation. The second benchmark system is ORC. The operating principle of ORC determines that all of the transmitting LEDs emit the same light intensity. In this way, a single point of a M-ary IM PAM constellation conveys binary information from the transmitter to the receiver. Note that there is a difference between the deployed IM PAM constellation of a space modulation technique and the corresponding PAM constellation of the benchmark systems. In the benchmark systems, a M-ary IM PAM constellation is constituted from the following set of points, $\mathcal{M}_{\mathrm{bs}} = \{s_0, \ldots, s_{M-1}\}$, where $s_0 = 0$ (zero light intensity). In contrast, as stated in Section II, the points of a PAM constellation deployed by a space modulation take only non-zero values. However, in order to enforce a fair comparison between all transmission schemes, the transmitted symbol vector $\mathbf{x}$ is normalized to the same average optical power. Finally, the detector of the benchmark systems is based on the ML principle. Thus, their detector is given from (6).

An indoor three dimensional space is considered where four transmitting LEDs and four receiving PDs are placed. The coordinates of the LEDs and PDs are given in Table I. The visual representation of the considered 4×4 ($N_t = 4$ and $N_r = 4$) VLC system is given in Fig. 1. The orientation of all the LEDs is given from the following orthonormal vector, $\mathbf{o}_t = [0, 0, -1]^T$, while the orientation of all the PDs is given as, $\mathbf{o}_r = [0, 0, 1]^T$. Furthermore, the transmitter semi-angle, $\Phi_{\frac{1}{2}}$, is 15 degrees. The area of each PD is 1 cm^2. The value of the responsivity of the PDs is $r = 0.4$ A/W. Finally, the FOV of the PDs, $\Psi_{\frac{1}{2}}$, is 30 degrees.



(a) BER versus γ_e (dB) of a 4×4 system with spectral efficiency 4 bpcu.



(b) BER versus γ_e (dB) of a 4×4 system with spectral efficiency 8 bpcu.

Fig. 2. Performance evaluation of the considered space modulation techniques (OSM, OGeSM, and OMS-SM) against the benchmark systems (OSMX and ORC). The system setup is 4×4 with spectral efficiency 4 and 8 bpcu. The solid lines correspond to simulation results, while the dashed lines correspond to the upper bounds derived in Section III.

Fig. 2(a) and 2(b) present the performance of the studied VLC system when the spectral efficiency is 4 bpcu and 8 bpcu, respectively. In each case, the spectral efficiency is set to the desired value by selecting the appropriate order of the employed IM PAM constellation. The metric of BER is plotted versus the *transmit* electrical SNR (as defined in Section III). Note that, due to the effect of the pathloss of the optical channel, the *detection* SNR at the side of the receiver faces a significant reduction with respect to the *transmit* electrical SNR. For this reason, the *transmit* electrical SNR (γ_e) in Fig. 2(a) and 2(b) takes high values.

As shown in Fig. 2(a) and 2(b), the analytical bounds of Section III are tight in relative high electrical SNR. In relative low electrical SNR, the theoretical bounds demonstrate a gap from the simulated curves. However, this phenomenon is a well known characteristic of the union bound technique deployed

in Section III [17]. Note that the theoretical ABEP of the benchmark systems is also evaluated following the framework of Section III. The is done by setting the appropriate values for $\mathcal{B}$ and k_t.

The inspection of Fig. 2(a) shows that the best BER performance in relative low SNRs ($\gamma_e < 110.5$ dB) is achieved by OSM. In contrast, as the value of SNR is increased above 110.5 dB, ORC has the best BER performance. The reason that ORC outperforms the other schemes is its operating principle, which resembles Single-Input Single-Output (SISO) communication. In SISO communication, only one symbol is conveyed from the transmitter to the receiver, just like ORC. In general, MIMO communication undergoes a performance degradation when the similarity between the existing sub-channels is high. Indeed, the channel similarity in the studied system setup is high due to the small spacing of the LEDs and PDs, and their symmetrical deployment. In fact, this is also the reason that OSM and OMS-SM outperform OSMX. The results in Fig. 2(a) show that OSM and OMS-SM are more robust compared to OSMX to channel similarity. Furthermore, Fig. 2(a) demonstrates that OSM outperforms OMS-SM for the same reason. In fact, OMS-SM is more prone to performance degradation due to channel similarity. This happens because OMS-SM spatially modulates multiple data streams. In relative high SNR, the worst performance is achieved by OGeSM.

Fig. 2(b) demonstrates that, even when the spectral efficiency is increased to 8 bpcu, OSM outperforms all the other schemes in relative low SNR ($\gamma_e < 129.8$ dB). However, at 129.8 dB, there is a crossing point after which OMS-SM becomes the most efficient transmission scheme. In fact, as shown in Fig. 2(b), for a BER=10^{-4}, OMS-SM exhibits an electrical SNR gain of about 5 dB compared to the OSMX. Fig. 2(b) shows that OSMX is the second most efficient scheme in relative high SNR. Furthermore, in relative high SNR, the performance of OSM, OGeSM, and ORC becomes worse compared to OSMX and OMS-SM. This is due to the higher order of the deployed constellation of OSM, OGeSM, and ORC. More specifically, OMS-SM, OSMX, OSM, OGeSM, and ORC use a constellation order of 4, 4, 64, 64, and 256, respectively. Therefore, it can be concluded that, in relative high SNR and for high spectral efficiency, it is preferable to exploit the multiplexing gain of OMS-SM and OSMX instead of the robustness of ORC and OSM to channel similarity. Finally, Fig. 2(b) shows that OGeSM gives the worst BER performance in values of SNR above 140 dB.

V. Conclusions

In this paper, the BER performance of OSM, OGeSM, and OMS-SM is studied against the corresponding performance of the benchmark systems of ORC and OSMX. The performance evaluation was conducted using both simulation and analytical results. As regard the theoretical results, tight upper bounds for the ABEP of all considered optical MIMO transmission schemes are provided. In this way, the provided simulation result are confirmed. For the studied system setup, it was concluded that OSM exhibits the best BER performance among the different schemes in relative low SNR. It was inferred that as SNR increases and the spectral efficiency is also increased, the performance of OMS-SM becomes the best one.

Acknowledgement

Prof. Harald Haas acknowledges support by EPSRC under grant EP/K008757/1.

References

[1] D. Tsonev, S. Videv, and H. Haas, "Light Fidelity (Li-Fi): Towards All-Optical Networking," in *Proc. SPIE, Broadband Access Commun. Technol. VIII*, vol. 9007, Dec. 18 2013. [Online]. Available: http://dx.doi.org/10.1117/12.2044649

[2] H. Chun, P. Manousiadis, S. Rajbhandari, D. Vithanage, G. Faulkner, D. Tsonev, J. McKendry, S. Videv, E. Xie, E. Gu, M. Dawson, H. Haas, G. Turnbull, I. Samuel, and D. O'Brien, "Visible Light Communication Using a Blue GaN LED and Fluorescent Polymer Color Converter," *IEEE Photon. Technol. Lett.*, vol. 26, no. 20, pp. 2035–2038, Oct. 2014.

[3] D. Tsonev, H. Chun, S. Rajbhandari, J. McKendry, S. Videv, E. Gu, M. Haji, S. Watson, A. Kelly, G. Faulkner, M. Dawson, H. Haas, and D. O'Brien, "A 3-Gb/s Single-LED OFDM-Based Wireless VLC Link Using a Gallium Nitride μLED," *IEEE Photon. Technol. Lett.*, vol. 26, no. 7, pp. 637–640, Apr. 2014.

[4] L. Zeng, D. O'Brien, H. Minh, G. Faulkner, K. Lee, D. Jung, Y. Oh, and E. T. Won, "High Data Rate Multiple Input Multiple Output (MIMO) Optical Wireless Communications Using White LED Lighting," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 9, pp. 1654–1662, Dec. 2009.

[5] G. J. Foschini and M. J. Gans, "On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas," *Wireless Personal Communications*, vol. 6, no. 6, pp. 311–335, 1998.

[6] J. M. Kahn and J. R. Barry, "Wireless Infrared Communications," *Proc. IEEE*, vol. 85, no. 2, pp. 265–298, Feb. 1997.

[7] T. Fath and H. Haas, "Performance Comparison of MIMO Techniques for Optical Wireless Communications in Indoor Environments," *IEEE Trans. on Commun.*, vol. 61, no. 2, pp. 733–742, February 2013.

[8] ——, "Optical Spatial Modulation Using Colour LEDs," in *Proc. of the IEEE 2013 Intern. Conf. on Commun. (ICC)*, June 2013, pp. 3938–3942.

[9] R. Mesleh, R. Mehmood, H. Elgala, and H. Haas, "Indoor MIMO Optical Wireless Communication Using Spatial Modulation," in *IEEE International Conference on Communications (ICC)*, Cape Town, South Africa, May 22–27 2010, pp. 1–5.

[10] Z. Yu, R. Baxley, and G. Zhou, "Multi-User MISO Broadcasting for Indoor Visible Light Communication," in *Proc. of the 2013 IEEE Intern. Conf. on Acoust., Speech and Signal Proc. (ICASSP)*, May 2013, pp. 4849–4853.

[11] K.-H. Park, Y.-C. Ko, and M. Alouini, "On the Power and Offset Allocation for Rate Adaptation of Spatial Multiplexing in Optical Wireless MIMO Channels," *IEEE Trans. on Commun.*, vol. 61, no. 4, pp. 1535–1543, April 2013.

[12] H. Ma, L. Lampe, and S. Hranilovic, "Robust MMSE Linear Precoding for Visible Light Communication Broadcasting Systems," in *Proc. of the 2013 IEEE Globecom Workshops*, Dec 2013, pp. 1081–1086.

[13] M. Di Renzo, H. Haas, A. Ghrayeb, S. Sugiura, and L. Hanzo, "Spatial Modulation for Generalized MIMO: Challenges, Opportunities, and Implementation," *Proc. IEEE*, vol. 102, no. 1, pp. 56–103, Jan 2014.

[14] A. Stavridis, S. Sinanović, M. D. Renzo., and H. Haas, "Energy Evaluation of Spatial Modulation at a Multi-Antenna Base Station," in *Proc. of the 78th IEEE Veh. Tech. Conf. (VTC)*, Las Vegas, USA, Sep. 2–5, 2013.

[15] A. Younis, N. Serafimovski, R. Mesleh, and H. Haas, "Generalised Spatial Modulation," in *Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, USA, Nov. 2010.

[16] J. Wang, S. Jia, and J. Song, "Generalised Spatial Modulation System with Multiple Active Transmit Antennas and Low Complexity Detection Scheme," *IEEE Trans. on Wireless Commun.*, vol. 11, no. 4, pp. 1605 – 1615, April 2012.

[17] J. G. Proakis and M. Salehi, *Communication System Engineering*. Prentice Hall, 1994.

[18] J. Barry, J. Kahn, W. Krause, E. Lee, and D. Messerschmitt, "Simulation of Multipath Impulse Response for Indoor Wireless Optical Channels," *IEEE J. Select. Areas Commun.*, vol. 11, no. 3, pp. 367–379, Apr. 1993.

[19] M. Chiani, D. Dardari, and M. K. Simon, "New Exponential Bounds and Approximations for the Computation of Error Probability in Fading Channels," *IEEE Trans. on Wireless Commun.*, vol. 2, no. 4, pp. 840–845, July 2003.

F.13 IEEE 2015 Global Commun. Conf. (GLOBECOM)

On the Performance of Multi-Stream Receive Spatial Modulation in the MIMO Broadcast Channel

Athanasios Stavridis[(1)], Marco Di Renzo[(2)], and Harald Haas[(1)]
[(1)]Li-Fi Research and Development Centre, Institute for Digital Communications (IDCOM),
School of Engineering, The University of Edinburgh, Edinburgh, UK
[(2)] Paris-Saclay University, Laboratory of Signals and Systems (UMR-8506),
CNRS - CentraleSupelec - University Paris-Sud XI, Paris, France
E–Mail: {a.stavridis, h.haas}@ed.ac.uk, marco.direnzo@lss.supelec.fr

***Abstract*—In this paper, a novel architecture for the Multiple-Input Multiple-Output (MIMO) broadcast channel is proposed and studied. The new architecture is based on the concept of Multi-Stream Receive-Spatial Modulation (MSR-SM). MSR-SM is a closed-loop transmission scheme, which applies the concept of multi-stream space modulation at the receiver side. A new and accurate framework for computing the Average Bit Error Probability (ABEP) of the new architecture is proposed. In addition, the new architecture is compared against state-of-the-art MIMO transmission in the broadcast channel and it is shown to: i) provide superior Bit Error Rate (BER) performance in the high Signal-to-Noise-Ratio (SNR) regime and ii) reduce the signal processing complexity at the transmitter.**

I. Introduction

Multiple-Input Multiple-Output (MIMO) communication has been considered as a promising approach for achieving high data rates without additional radio resources. This is the direct result of the information-theoretic analysis of MIMO communication [1, 2]. In addition, multi-antenna architectures are able to obtain a Spatial MultipleXing (SMX), diversity, or Signal-to-Noise Ratio (SNR) gain, depending on the deployed system setup. Unfortunately, in MIMO communication, as the system size is increased, its complexity, in terms of signal processing and hardware, is also increased. In fact, this problem becomes even more challenging in massive MIMO deployments.

A MIMO concept that has the potential to provide a low complexity system implementation is Spatial Modulation (SM) [3–5]. The operating principle of SM is designed in such a way that the transmitter requires only one Radio Frequency (RF) chain. Therefore, significant Energy Efficiency (EE) and complexity gains are obtained in comparison to conventional MIMO techniques [6]. Due to the large potential of SM, several variants of SM can be found in the published research [7–9]. A comprehensive overview of the existing literature on SM is given in [3].

As demonstrated in [3], the study of SM and its variants in point-to-point communication is extensive. In addition, the formation of a Multi-User (MU) system based on the concept of SM can be undertaken with the aid of a multiple access scheme, such as Time Division Multiple Access (TDMA), Frequency Division Multiple Access (FDMA), or Orthogonal Frequency-Division Multiple Access (OFDMA). However, a new trend in wireless communication promotes the aggressive allocation of multiple users in the same time and frequency resources. In such systems, usually, the inherent interference is eliminated or mitigated via the deployment of Space Division Multiple Access (SDMA) techniques. In particular, the Multiple Access Channel (MAC) is formed in the uplink [10] and the broadcast channel is established in the downlink [11].

Relevant to this field, [12–15] extend the concept of SM in a MU setup. In particular, the schemes proposed in [12, 13] are applicable to the uplink, whereas the scheme in [14, 15] is suitable for the downlink. Unfortunately, in [12, 13], either there is an increase in the the complexity of the receiver, or there is an error saturation in the Bit Error Rate (BER) performance. In addition, the scheme in [14, 15] experiences a BER performance degradation with respect to the Single-User (SU) communication. Note that the incorporation of SM in a SDMA architecture is a challenging task since the deployment of interference elimination or reduction techniques is difficult. Especially in the downlink, this difficulty originates from the way in which information is transmitted in SM. Thus, the proposition of a SDMA-based MU architecture for SM becomes a challenging and important task.

Against this background and based on the concept of Multi-Stream Receive-Spatial Modulation (MSR-SM), this paper proposes and studies a new MU architecture for the downlink. MSR-SM is a closed loop and point-to-point modulation scheme which applies the concept of Multi-Stream Spatial Modulation (MS-SM) at the receiver side [16, 17]. In more detail, this paper incorporates MSR-SM in the MIMO broadcast channel. This is achieved using Zero Forcing (ZF) forcing precoding.

It is demonstrated that the new architectures outperforms the corresponding conventional spatially multiplexed MIMO broadcast channel, in terms of BER, in high SNR. In addition, this paper derives tight upper bounds for the theoretical Average Bit Error Probability (ABEP) of a typical user when: i) the wireless channel follows a Rayleigh distribution and ii) Perfect-Channel State Information at the Transmitter (P-CSIT) is available. This is undertaken by considering an accurate statistical framework for the received signal. The theoretical ABEP of point-to-point MSR-SM has been recently studied in [18, 19]. However, it is emphasized that the mathematical framework derived in this paper exhibits major differences with respect to those presented in [18, 19]. In [18], the

statistical description of the received signal is not taken into account. Furthermore, the ABEP computed in [19] is derived for a scenario where a suboptimal detector, which decouples the detection process, is deployed. In contrast, the analysis given in this paper is different for the following reasons: i) multiple users are considered; ii) the statistical description of the received signal of a typical user is considered; and iii) the detection process is based on the Maximum Likelihood (ML) principle, which imposes some additional mathematical difficulty. Note that the new framework is directly applicable to a point-to-point single user scenario.

The rest of the paper is organized as follows: The system model of the new architecture is introduced in Section II. The computational complexity of the new architecture is discussed in Section III. The theoretical analysis of the ABEP of a typical user is presented in Section IV. The proposed MU MSR-SM architecture is compared against the corresponding benchmark system in Section V. In addition, numerical results that validate the new theoretical findings are presented in the same section. Finally, the concluding comments of this paper are given in Section VI.

Notation: In the following, lowercase bold letters denote vectors and uppercase bolt letters denote matrices. $(\cdot)^T$, $(\cdot)^H$, tr$(\cdot)$ and $\mathbf{A}^{1/2}$ denote transpose, Hermitian transpose, matrix trace and the square root of $\mathbf{A}$, respectively. $\|\cdot\|_2$ represents the Euclidean norm. $\mathbf{diag}\left(a_1,\ldots,a_n\right)$ represents a diagonal matrix whose main diagonal includes the elements $a_1,\cdots,a_n$. E$[\cdot]$ is the mean value of a RV. A complex Gaussian distribution with mean m and variance σ_C^2 is represented as $\mathcal{CN}(m,\sigma_C^2)$, where its real and imaginary part are independent and identically distributed (i.i.d.) Gaussian RV with distribution $\mathcal{N}(m,\frac{\sigma_C^2}{2})$. Re$\{\cdot\}$ denotes the real part of a complex number or matrix.

II. System Model

In this section, a single cell uncoded downlink transmission is considered. More specifically, a Base Station (BS) with N_t antennas aims to establish the MIMO broadcast channel in order to serve N_u users. Each user is equipped with N_r antennas. Since the transmitter is a BS, the assumption that $N_t \geq N_u N_r$ is realistic. Furthermore, the MIMO wireless channel between the transmitter and the remote users is assumed to be frequency flat and quasi-static. Finally, in this paper, a scenario where the transmitter possesses P-CSIT, using either the channel reciprocity or fast and error free links from the users, is considered[1].

Due to the availability of P-CSIT, linear precoding can be applied at the transmitter. In addition, the N_t transmit antennas and the $B = N_u N_r$ receive antennas can be interpreted as a $N_t \times B$ MIMO system. In this case, the matrix form of the system equation is given as:

$$\mathbf{y} = \mathbf{HPDx} + \mathbf{w}, \tag{1}$$

[1]Note that supplying the transmitter with P-CSIT is difficult task. Usually, the transmitter is supplied with Imperfect-Channel State Information at the Transmitter (I-CSIT). However, the study of the effect of I-CSIT is out of the scope of this paper.

where, $\mathbf{y}$ is a $N_u N_r \times 1$ vector which is written as:

$$\mathbf{y} = \left[\mathbf{y}_1^T,\ldots,\mathbf{y}_{N_u}^T\right]^T. \tag{2}$$

Here, $\mathbf{y}_i$, $i=1,\ldots,N_u$, is the $N_r \times 1$ received signal vector at the i-th user. In addition, the $N_u N_r \times N_t$ wireless channel is denoted as:

$$\mathbf{H} = \left[\mathbf{H}_1^H,\ldots,\mathbf{H}_{N_u}^H\right]^H, \tag{3}$$

where, the sub-matrix $\mathbf{H}_i$, $i=1,\ldots,N_u$, denotes the channel from the transmitter to the i-th user. Furthermore, due to rich scattering, no channel correlation is assumed. In this paper, large scale fading is not considered. Therefore, it is assumed that each sub-matrix $\mathbf{H}_i$ is distributed as $\mathbf{H}_i \sim \mathcal{CN}(\mathbf{0},\mathbf{I})$. In (1),

$$\mathbf{P} = [\mathbf{P}_1,\ldots,\mathbf{P}_{N_u}] \tag{4}$$

denotes the $N_t \times N_u N_r$ linear precoding matrix. Here, $\mathbf{P}_i$, $i=1,\ldots,N_u$, is the corresponding precoding matrix for the i-th user. In order to enforce a constrained power transmission, a diagonal normalization matrix:

$$\mathbf{D} = \mathbf{diag}\left(d_1,\ldots,d_{N_u N_r}\right) \tag{5}$$

is deployed. Here, each element d_i, $i=1,\ldots,N_u N_r$, of the main diagonal of $\mathbf{D}$ equals to:

$$d_i = \sqrt{\frac{1}{\|\mathbf{p}_i\|_2^2}}, \tag{6}$$

where, $\mathbf{p}_i$ corresponds to the i-th column of $\mathbf{P}$. In this way, every column of the normalized precoding matrix:

$$\mathbf{P}_{\text{norm}} = \mathbf{PD} \tag{7}$$

has unity power. Furthermore,

$$\mathbf{x} = \left[\mathbf{x}_1^T,\ldots,\mathbf{x}_{N_u}^T\right]^T \tag{8}$$

represents the collective information carrying symbol vector at the transmitter. Here, $\mathbf{x}_i$, $i=1,\ldots,N_u$, is the signal vector which carries binary information to the i-the user. Finally, the white Gaussian noise is denoted by:

$$\mathbf{w} = \left[\mathbf{w}_1^T,\ldots,\mathbf{w}_{N_u}^T\right]^T, \tag{9}$$

where, $\mathbf{w} \sim \mathcal{CN}(\mathbf{0},\sigma_{\mathbf{w}}^2\mathbf{I})$. Furthermore, $\mathbf{w}_i$, $i=1,\ldots,N_u$, is the Gaussian noise observed by the i-th user.

In this paper, the linear precoder is designed based on the ZF principle. The selection of ZF precoding is justified by its low complexity and its ability to totally eliminate the interference between different users and between different antennas of the same user. Therefore, ZF precoding can be considered as an efficient method for the formation of a MU architecture based on MSR-SM.

Given that the ZF precoder is the pseudo-inverse of the channel matrix $\mathbf{H}$, its matrix form is written as:

$$\mathbf{P} = \mathbf{H}^H\left(\mathbf{HH}^H\right)^{-1}. \tag{10}$$

In this case, the i-th element of the main diagonal of $\mathbf{D}$ equals

$$\mathbf{x}_i = \left[0,\ldots,0,\ \underbrace{s_1}_{i_1\text{-th position}},0,\ldots,0,\ \underbrace{s_i}_{i_k\text{-th position}},0,\ldots,0,\ \underbrace{s_{N_s}}_{i_{N_s}\text{-th position}}\ 0\ldots,0\right]^T \tag{14}$$

to:

$$d_i = \sqrt{\frac{1}{\left[\left(\mathbf{H}\mathbf{H}^H\right)^{-1}\right]_{i,i}}},\ i = 1,\ldots,N_uN_r. \tag{11}$$

The incorporation of (10) in (1) gives:

$$\mathbf{y} = \mathbf{D}\mathbf{x} + \mathbf{w}. \tag{12}$$

In addition, the received signal at the i-th is given as:

$$\mathbf{y}_i = \mathbf{D}_i\mathbf{x}_i + \mathbf{w}_i,\quad i = 1,\ldots,N_u. \tag{13}$$

In (13), $\mathbf{D}_i$, $i = 1,\ldots,N_u$, is the $N_r \times N_r$ diagonal normalization matrix of the corresponding precoding matrix $\mathbf{P}_i$. In particular, the normalization matrix $\mathbf{D}$ can be interpreted as the following block diagonal matrix, $\mathbf{D} = \mathbf{diag}\left(\mathbf{D}_1,\ldots,\mathbf{D}_{N_u}\right)$.

From (13), it can be inferred that the structure of $\mathbf{x}_i$, $i = 1,\ldots,N_u$, determines the way that information is transmitted to each user. For example, conventional SMX MIMO transmission is established if all of the elements of $\mathbf{x}_i$ are drawn from a conventional M-ary constellation diagram $\mathcal{M}$.

In order to establish a MSR-SM transmission mechanism between the transmitter and the i-th user, every symbol period, the following two requirements have to be fulfilled. Firstly, $N_s \leq N_r$ conventional symbols have to conveyed from the transmitter to the the i-th user per symbol period. In this way, only a subset of N_s antennas at the i-th user receive a non-zero signal. All of the other antennas face only thermal noise. And secondly, additional binary information has to be encoded via the indices of the N_s (out of N_r) receiving antennas. Note that when $N_s = N_r$, MSR-SM reduces to conventional SMX transmission and no additional information is conveyed via the indices of the receiving antennas.

Provided that the deployed precoding method is ZF, the received signal at each user is given in (13). Hence, via the appropriate selection of the structure of $\mathbf{x}_i$, the transmitter is able to impose that the noise free received signal $\mathbf{D}_i\mathbf{x}_i$ has exactly N_s non-zero elements and $N_r - N_s$ zero elements. Therefore, the non-zero elements of $\mathbf{D}_i\mathbf{x}_i$ is a scaled version of the corresponding non-zero elements of $\mathbf{x}_i$. Also, the positions of the zero elements of $\mathbf{D}_i\mathbf{x}_i$ are the same as the positions of the zero elements of $\mathbf{x}_i$. Thus, a number of bits can be encoded on the positions of the non-zero elements of $\mathbf{x}_i$.

The structure of the transmission alphabet of MSR-SM, $\mathcal{B}_i$, is given in (14) at the top of this page. As shown in (14), $\mathbf{x}_i \in \mathcal{B}_i$ has exactly N_s non-zero elements which belong to a conventional constellation, $\{s_1,\ldots,s_{N_s}\} \in \mathcal{M}$. Here, $\mathcal{M}$ denotes the deployed constellation. The positions of the non-zero elements of $\mathbf{x}_i$ correspond to the indices of the signal receiving antennas, while the zero elements of $\mathbf{x}_i$ correspond to the indices of the non-receiving antennas. Given that the length of $\mathbf{x}_i$ is N_r and there are N_s non zero elements, the number of total combinations of N_s non-zero elements (receiving antennas) out of N_r is $\binom{N_r}{N_s}$. Here, $\binom{\cdot}{\cdot}$ denotes the binomial coefficient. However, only $k_c = 2^{\log_2\left(\lfloor\binom{N_r}{N_s}\rfloor\right)}$ combinations are used in order to encode $k_2^{\text{MSR-SM}} = \log_2\left(\lfloor\binom{N_r}{N_s}\rfloor\right)$ bits. This is done by assigning a unique binary index of length of $k_2^{\text{MSR-SM}}$ bits to each one of the k_c used (legal) combinations. The selection of these (legal) combinations of receiving antennas can be done intelligently in order to minimize the instantaneous BER or it can be done randomly. In the first case, the system complexity is increased and the selection is undertaken adaptively. Given that MSR-SM is a closed loop scheme, the latter case is expected to offer a good performance with no further complexity overhead. In this paper, the focus is on the latter case.

In this way, every symbol period, the bit-stream to be transmitted is divided in two parts. The first part of $k_1^{\text{MSR-SM}} = N_s\log_2(M)$ bits, is encoded and transmitted via N_s symbols drawn from the M-ary constellation $\mathcal{M}$. The second part of $k_2^{\text{MSR-SM}} = \log_2\left(\lfloor\binom{N_r}{N_s}\rfloor\right)$ bits is encoded on the combination of the signal receiving antennas. Therefore, the spectral efficiency of MSR-SM is:

$$k_{\text{MSR-SM}} = N_s\log_2(M) + \log_2\left(\lfloor\binom{N_r}{N_s}\rfloor\right) \tag{15}$$

bits per channel use (bpcu) per user.

Assuming that the i-th user is aware of $\mathbf{D}_i$, the reconstruction of the transmitted bit-streams is conducted by detecting the transmitted vectors $\mathbf{x}_i$, $i = 1,\ldots,N_u$. This is undertaken at every user independently by using the following ML detector:

$$(\tilde{\mathbf{x}}_i) = \arg\min_{\mathbf{x}_i}\|\mathbf{y}_i - \mathbf{D}_i\mathbf{x}_i\|_2^2,\quad i = 1,\ldots,N_u. \tag{16}$$

III. Computation Complexity Analysis

The inspection of (14) shows that the sparse structure of the transmission alphabet of MSR-SM, $\mathcal{B}_i$, can be deployed in order to decrease the computational complexity of the transmitter. Assuming that the ZF precoder in (10) is precomputed before the transmission of each block of symbols, the transmitted signal, $\mathbf{s} = \mathbf{P}\mathbf{D}\mathbf{x}$, in (1) can be calculated with $\mathcal{C}_t = N_t(8N_uN_s - 2) + 2N_uN_s$ real operations (additions or multiplications). Hence, lower values of N_s result in lower computational complexity at the transmitter. Note that when it holds that $N_s = N_t$, MSR-SM is transformed into a conventional SMX transmission. In this way, it can be inferred that, at the transmitter side, the computational complexity of MSR-SM is less than the corresponding complexity of conventional SMX transmission.

At the users' side, the ML detection process of the conventional MIMO broadcast channel is decoupled in per

$$\delta_{k+1} = \begin{cases} 1, & k = -1, \\ \frac{k}{k+1} \sum_{i=1}^{k+1} \left[\sum_{j=1}^{N} \left(1 - \frac{\alpha_1}{\alpha_j} \right)^i \right] \delta_{k+1-i}, & k = 0, 1, 2, \ldots . \end{cases} \quad (25)$$

single stream detection. Therefore, its detection complexity is lower compared with the corresponding complexity of MU MSR-SM. However, in [19] a suboptimal detector for MSR-SM is proposed. This detector achieves almost the same computational complexity as the previous decoupled detector. The study of the detector of [19] is, however, outside of the scope of this paper.

IV. Theoretical Evaluation of the Average Bit Error Probability

In this section, the ABEP of a typical user is derived. Based on the union bound technique, the ABEP of the i-th user, $P_{\text{bit}}^{i}(\gamma)$, for a given transmit SNR γ, is bounded as:

$$P_{\text{bit}}^{i}(\gamma) \leq \frac{1}{|\mathcal{B}_i| k_{\text{MSR-SM}}} \sum_{\mathbf{x}_i} \sum_{\substack{\hat{\mathbf{x}}_i \\ \hat{\mathbf{x}}_i \neq \mathbf{x}_i}} d(\mathbf{x}_i \rightarrow \hat{\mathbf{x}}_i) P_{\text{e}}^{i}(\mathbf{x}_i \rightarrow \hat{\mathbf{x}}_i, \gamma). \quad (17)$$

In (17), $P_{\text{e}}^{i}(\mathbf{x}_i \rightarrow \hat{\mathbf{x}}_i, \gamma)$ denotes the Pairwise Error Probability (PEP) of transmitting $\mathbf{x}_i$ to the i-th user while its detector erroneously decides in favor of $\hat{\mathbf{x}}_i$. The Hamming distance between the bit-words represented by $\mathbf{x}_i$ and $\hat{\mathbf{x}}_i$ is denoted as $d(\mathbf{x}_i \rightarrow \hat{\mathbf{x}}_i)$. Also, the size of the transmission alphabet of MSR-SM to the i-th user is given as $|\mathcal{B}_i| = M^{N_s} 2^{\lfloor \log_2 \binom{N_r}{N_s} \rfloor}$.

The evaluation of (17) requires the knowledge of $P_{\text{e}}^{i}(\mathbf{x}_i \rightarrow \hat{\mathbf{x}}_i, \gamma)$, which is the expectation of the instantaneous PEP over all channel realizations. The instantaneous PEP of the i-th user is expressed as:

$$P_{\text{e}}^{i}\left(\mathbf{x}_i \rightarrow \hat{\mathbf{x}}_i, \gamma | \mathbf{D}_i^2\right) = Q\left(\sqrt{\frac{\mathbf{c}_i^H \mathbf{D}_i^2 \mathbf{c}_i}{2} \gamma} \right) = Q\left(\sqrt{\frac{z_i \gamma}{2}} \right), \quad (18)$$

where, $\mathbf{c}_i = \mathbf{x}_i - \hat{\mathbf{x}}_i$; $z_i = \mathbf{c}_i^H \mathbf{D}_i^2 \mathbf{c}_i$; and $\gamma = 1/\sigma_{\mathbf{w}_i}^2$ is the transmit SNR. The proof of (18) relies on the fact that a symbol error occurs at the i-th user when

$$\mathcal{E}_i\left(\mathbf{x}_i, \hat{\mathbf{x}}_i\right) = \left\{ \|\mathbf{y}_i - \mathbf{D}_i \mathbf{x}_i\|_2^2 > \|\mathbf{y}_i - \mathbf{D}_i \hat{\mathbf{x}}_i\|_2^2 \right\}. \quad (19)$$

From (19), it is not difficult to show that:

$$\mathcal{E}_i\left(\mathbf{x}_i, \hat{\mathbf{x}}_i\right) = \left\{ -\text{Re}\{\mathbf{c}_i^H \mathbf{D}_i \mathbf{w}_i\} > \frac{\mathbf{c}_i^H \mathbf{D}_i^2 \mathbf{c}_i}{2} \right\}. \quad (20)$$

Therefore, using the statistical distribution of $\mathbf{w}_i$, the PEP of the i-th user is expressed as in (18).

The inspection of the instantaneous PEP of the i-th user in (18) shows that it is conditioned on the Random Variable (RV) z_i. Thus, the Probability Density Function (PDF) of z_i has to be derived. However, it holds that:

$$z_i = \sum_{k=1}^{N_r} |x_k - \hat{x}_k|^2 d_k^2 = \sum_{x_k - \hat{x}_k \neq 0} |x_k - \hat{x}_k|^2 d_k^2. \quad (21)$$

In [20], it is explicitly shown that d_k is a gamma RV with $d_k^2 \sim \text{Gamma}\left(L_{\text{MU}}, 1\right)$, where, $L_{\text{MU}} = N_t - N_u N_r + 1$. Therefore, the RV $Z_k = |x_k - \hat{x}_k|^2 d_k^2$ is also a gamma RV, when it holds that $x_k - \hat{x}_k \neq 0$. Here, for notational convenience, the following variable is introduced, $b_k = |x_k - \hat{x}_k|^2$. Based on the previous arguments, the PDF of Z_k is given as:

$$f_{Z_k}(x) = \frac{1}{b_k^L \Gamma(L)} x^{L-1} e^{\frac{x}{b_k}} H_0(x), \quad (22)$$

where, $H_0(x)$ is the Heaviside step function, for which it holds that: $H_0(x) = 0$ for $x < 0$ and $H_0(x) = 1$ for $x \geq 0$.

Usually, in published research, the RVs d_k^2 are assumed to be statistically independent in order to simplify the whole analysis [21, 22]. This assumption is in contradiction with the structure of $d_k^2 = 1/\left[\left(\mathbf{H}_i \mathbf{H}_i^H \right)^{-1} \right]_{k,k}$, since the realization of every RV d_k^2 occurs using the same mathematical operation on the same random matrix $\mathbf{H}_i$.

In this paper, the statistical dependence of the RVs d_k^2 is considered. For this reason, z_i is the result of the sum of correlated gamma RVs. Thus, using the result form [23, Corollary 1] and (21), the PDF of z_i is directly expressed as:

$$f_{z_i}(x) = \left[\prod_{l=1}^{N_i} \left(\frac{\alpha_1}{\alpha_l} \right)^L \right] \left[\sum_{k=0}^{+\infty} \frac{\delta_k x^{N_i L + k - 1} e^{-\frac{x}{\alpha_1}}}{\alpha_1^{N_i L + k} \Gamma\left(N_i + k\right)} \right] H_0(x), \quad (23)$$

where, N_i is the number of non zero elements of $\mathbf{c}_i$ for a given pair of $\mathbf{x}_i$ and $\hat{\mathbf{x}}_i$. In (23), α_l, $l = 1, \ldots, N_i$, denote the eigenvalues of $\mathbf{A} = \mathbf{B}\mathbf{R}$ in ascending order. Here, $\mathbf{B}$ is defined as the following diagonal matrix, $\mathbf{B} = \mathbf{diag}\left(b_1, \ldots, b_{N_i}\right)$, where b_l, $l = 1, \ldots, N_i$ is the absolute value of the l-th non zero element of $\mathbf{c}_i$. Also, $\mathbf{R}$ is a $N_i \times N_i$ matrix defined as:

$$\mathbf{R} = \begin{bmatrix} 1 & \sqrt{\rho_c} & \cdots & \sqrt{\rho_c} \\ \sqrt{\rho_c} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \sqrt{\rho_c} \\ \sqrt{\rho_c} & \cdots & \sqrt{\rho_c} & 1 \end{bmatrix}, \quad (24)$$

where, ρ_c is the Pearson product-moment correlation coefficient between any pair of two different RVs of the main diagonal of $\mathbf{D}_i^2$. Finally, δ_k, $k = 0, 1, 2, \ldots$, are given in (25) at the top of this page.

In order to further simplify (18), the Chernoff bound of the Q-function, $Q\left(x\right) \leq \frac{1}{2} e^{-\frac{x^2}{2}}$, is considered. In this way, the PEP of the i-th user is expressed as:

$$\begin{aligned} P_{\text{e}}^{i}\left(\mathbf{x}_i \rightarrow \hat{\mathbf{x}}_i, \gamma_i\right) &\leq \frac{1}{2} \text{E}_{z_i}\left[e^{-\frac{z_i \gamma_i}{4}} \right] \\ &= \frac{1}{2} \int_{-\infty}^{+\infty} e^{-\frac{y \gamma_i}{4}} f_{z_i}(y) \text{d}y. \end{aligned} \quad (26)$$

Provided that the PDF of z_i is given from (23), after some algebraic manipulations which are omitted due to space

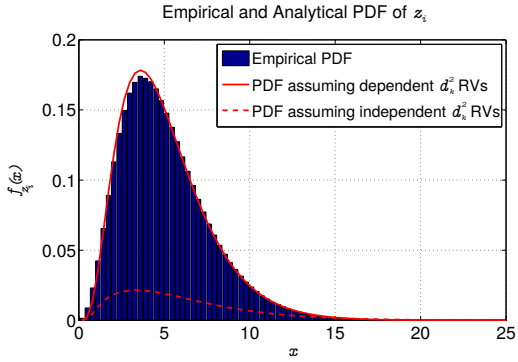


Fig. 1. Illustration of empirical and analytical PDF of (21) by assuming that: i) the RVs d_k, $k = 1, \ldots, 2$ are statistically dependent and ii) they are independent. Setup: $\mathbf{H} \sim \mathcal{CN}(\mathbf{0}_{2\times 4}, \mathbf{I}_{2\times 4})$; and ii) $b_1 = 0.5$ and $b_1 = 1.2$.

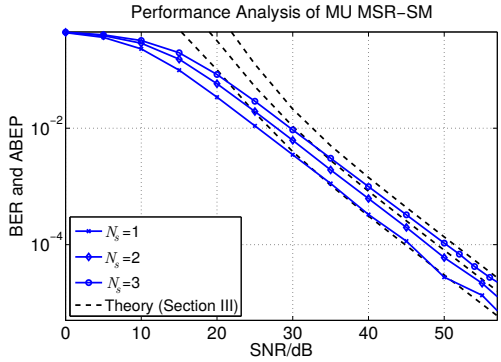


Fig. 2. Performance analysis of a typical user, when MSR-SM is deployed and the precoding method is ZF with P-CSIT: simulation results vs. the bounds in Section IV. Setup: $N_t = 16$, $N_r = 4$, $N_u = 4$, $N_s = \{1, 2, 3\}$, and $M = 4$.

limitation, the evaluation of (26) gives:

$$P_{\mathrm{e}}^{i}(\mathbf{x}_i \to \hat{\mathbf{x}}_i, \gamma) \leq \frac{\left[\prod_{l=1}^{N_i} \left(\frac{\alpha_1}{\alpha_l}\right)^{L_{\mathrm{MU}}}\right]}{2} \left(\frac{\alpha_1}{4}\xi_i\gamma + 1\right)^{-N_i L_{\mathrm{MU}}} \times \sum_{k=0}^{+\infty} \delta_k \left(\frac{\alpha_1}{4}\xi_i\gamma + 1\right)^{-k}. \quad (27)$$

Therefore, the computation of the ABEP of the i-th user is conducted from (17), by using (27). Note that the derived bounds are directly applicable to the conventional MIMO broadcast channel when $N_s = N_r$.

V. Simulation Results and Discussion

The main objectives of this section are: i) the validation of the theoretical results of Section IV and ii) the comparison of the new MU architecture with the corresponding State-of-the-Art (SotA) benchmark scheme. In this section, the wireless channel is assumed to follow a Raleigh distribution.

Section IV provides the derivation of the ABEP of MU MSR-SM using the PDF of z_i given in (23). The deployment of (23) relies on the fact that the RVs d_k^2, $k = 1, \ldots, N_r$, are statistically dependent. In order to confirm that the RVs d_k^2 are statistical dependent, Fig. 1 depicts the empirical PDF of (21) against its analytical form as given in (23). In addition, Fig. 1 presents the analytical PDF of (23) under the assumption that d_k^2, $k = 1, \ldots, N_r$, are statistically independent RVs. If this assumption was valid, the PDF of (21) could be directly obtained using the result from [23, Theorem 1]. From Fig. 1, it can be concluded that the theoretical PDF of (23) perfectly matches its empirical PDF. In contrast, when the RVs d_k^2, $k = 1, \ldots, N_r$ are assumed to be independent, the obtained PDF from [23, Theorem 1] deviates from the empirical results.

As a form of validation of the bounds derived in Section IV, Fig. 2 compares theses bounds against the corresponding BER curves obtained via Monte Carlo simulations. The inspection of Fig. 2 shows that the upper bounds are tight in high SNR regime. In low SNR, there is a gap between the theoretical and simulated curves. However, this gap is a well known effect of the union bound technique deployed in Section IV [24].

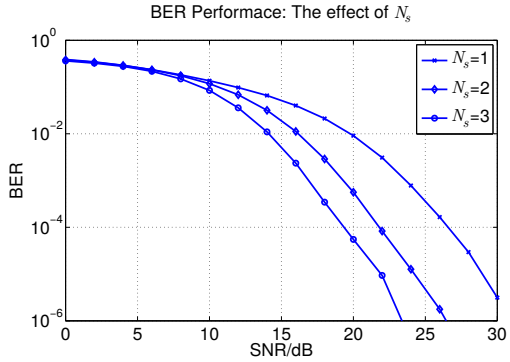


Fig. 3. BER performance of a typical user of MU MSR-SM as a function of N_s. Setup: $N_t = 20$, $N_r = 4$, $N_u = 4$. The spectral efficiency is 8 bpcu.

In Fig. 3, the BER performance of MU MSR-SM as a function of N_s is shown for the same spectral efficiency. In particular, Fig. 3 demonstrates that the BER of a typical user improves as N_s increases. This happens because higher values of N_s require a lower modulation order of Quadrature Amplitude Modulation (QAM) in order to achieve the same spectral efficiency. However, as shown in Fig 4, the optimal value of N_s is not to N_r, i.e., the conventional MIMO broadcast channel.

A comparison between the BER performance of MU MSR-SM and the corresponding benchmark system is presented in Fig. 4. In this paper, the selected benchmark system is the conventional spatially multiplexed MIMO broadcast channel, where N_r symbol streams are established per user. The spectral efficiency of both schemes is set to be the same by selecting the appropriate constellation order. In addition, MU MSR-SM spatially modulates $N_s = 2$ and $N_s = 3$

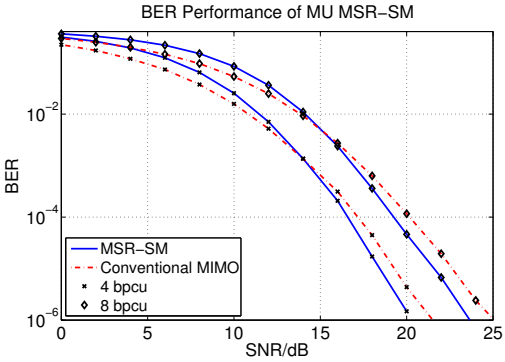


Fig. 4. BER performance of MSR-SM versus the conventional MIMO broadcast channel. Setup: $N_t = 20$, $N_r = 4$, and $N_u = 4$.

parallel symbol streams in order to achieve 4 and 8 bpcu, respectively. As shown in Fig. 4, in low SNR the conventional MIMO broadcast channel offers a slightly better performance than MU MSR-SM. However, for practical values of BER approximately less than 10^{-2}, MU MSR-SM outperforms the benchmark systems. The performance gap between the considered schemes is about 0.5 dB and 1 dB for 4 and 8 bpcu respectively. Furthermore, Fig. 4 indicates that MU MSR-SM achieves the same diversity order as the benchmark system but higher coding gain. The theoretical proof of the previous indication is a subject of our future work.

VI. Conclusions

The incorporation of MSR-SM in the MIMO broadcast channel was introduced and studied in this paper. In particular, a novel upper bound for the theoretical BER of the new architecture was derived. This was undertaken by using a new and accurate statistical framework for the received signal of a typical user. It was shown that the new bounds are tight. For the purpose of comparison, MU MSR-SM was evaluated, in terms of BER, against the corresponding conventional MIMO broadcast channel. It was shown that the new architecture outperforms the benchmark system in high SNR regime. Also, it offers lower computational complexity in the transmitter. Therefore, MSR-SM could be considered as an alternative transmission mechanism for the MIMO broadcast channel.

Finally, the derivation of the diversity order and coding of MU MSR-SM, as well as its theoretical comparison with the SotA is a subject of our future work.

Acknowledgement

Professor Harald Haas acknowledges support by EPSRC under grant EP/K008757/1.

References

[1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 6, pp. 311–335, 1998.

[2] E. Telatar, "Capacity of multi-antenna gaussian channels," *European Trans. on Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov. / Dec. 1999.

[3] M. Di Renzo, H. Haas, A. Ghrayeb, S. Sugiura, and L. Hanzo, "Spatial modulation for generalized MIMO: Challenges, opportunities, and implementation," *Proc. IEEE*, vol. 102, no. 1, pp. 56–103, Jan 2014.

[4] R. Mesleh, H. Haas, S. Sinanović, C. W. Ahn, and S. Yun, "Spatial modulation," *IEEE Trans. on Veh. Tech.*, vol. 57, no. 4, pp. 2228 – 2241, Jul. 2008.

[5] P. Yang, M. Di Renzo, Y. Xiao, S. Li, and L. Hanzo, "Design guidelines for spatial modulation," *IEEE Commun. Surveys Tut.*, vol. 17, no. 1, pp. 6–26, First Quarter 2015.

[6] A. Stavridis, S. Sinanović, M. D. Renzo., and H. Haas, "Energy evaluation of spatial modulation at a multi-antenna base station," in *Proc. of the 78th IEEE Veh. Tech. Conf. (VTC)*, Las Vegas, USA, Sep. 2–5, 2013.

[7] J. Jeganathan, A. Ghrayeb, L. Szczecinski, and A. Ceron, "Space shift keying modulation for MIMO channels," *IEEE Trans. on Wireless Commun.*, vol. 8, no. 7, pp. 3692–3703, Jul. 2009.

[8] M. Wen, X. Cheng, H. Poor, and B. Jiao, "Use of SSK modulation in two-way amplify-and-forward relaying," *IEEE Trans. on Veh. Techn.*, vol. 63, no. 3, pp. 1498–1504, March 2014.

[9] P. Som and A. Chockalingam, "Performance analysis of space-shift keying in decode-and-forward multihop MIMO networks," *IEEE Trans. on Veh. Techn.*, vol. 64, no. 1, pp. 132–146, Jan 2015.

[10] M. Matthaiou, N. Chatzidiamantis, G. Karagiannidis, and J. Nossek, "ZF detectors over correlated K fading MIMO channels," *IEEE Trans. on Commu.*, vol. 59, no. 6, pp. 1591–1603, June 2011.

[11] M. Kountouris and J. Andrews, "Downlink SDMA with limited feedback in interference-limited wireless networks," *IEEE Trans. on Wireless Commun.*, vol. 11, no. 8, pp. 2730–2741, August 2012.

[12] N. Serafimovski, S. Sinanovic, M. Renzo, and H. Haas, "Multiple access spatial modulation," *EURASIP J. on Wireless Commun. and Netw.*, vol. 2012, no. 1, pp. 1–20, 2012.

[13] M. Di Renzo and H. Haas, "Bit error probability of space-shift keying MIMO over multiple-access independent fading channels," *IEEE Trans on Veh. Technol.*, vol. 60, no. 8, pp. 3694–3711, Oct 2011.

[14] S. Narayanan, M. J. Chaudhary, A. Stavridis, R. Di Renzo, F. Graziosi, and H. Haas, "Multi-user spatial modulation MIMO," in *Proc. of IEEE Wireless Communications and Networking Conference (WCNC)*, Apr. 6–9, 2014.

[15] X. Li, Y. Zhang, L. Xiao, X. Xu, and J. Wang, "A novel precoding scheme for downlink multi-user spatial modulation system," in *Proc. IEEE 24th Int. Symp. Pers. Indoor and Mobile Radio Commun. (PIMRC)*, Sept 2013, pp. 1361–1365.

[16] J. Wang, S. Jia, and J. Song, "Generalised spatial modulation system with multiple active transmit antennas and low complexity detection scheme," *IEEE Trans. on Wireless Commun.*, vol. 11, no. 4, pp. 1605 – 1615, April 2012.

[17] K. Ntontin, M. Di Renzo, A. Perez-Neira, and C. Verikoukis, "Performance analysis of multistream spatial modulation with maximum-likelihood detection," in *Proc. of 2013 IEEE Global Commun. Conf. (GLOBECOM)*, Dec 2013, pp. 1590–1594.

[18] R. Zhang, L.-L. Yang, and L. Hanzo, "Generalised pre-coding aided spatial modulation," *IEEE Trans. on Wireless Commun.*, vol. 12, no. 11, pp. 5434–5443, November 2013.

[19] ——, "Error probability and capacity analysis of generalised pre-coding aided spatial modulation," *IEEE Trans. on Wireless Commun.*, vol. PP, no. 99, pp. 1–1, 2014.

[20] D. Gore, R. Heath, and A. Paulraj, "Transmit selection in spatial multiplexing systems," *IEEE Commun. Lett.*, vol. 6, no. 11, pp. 491–493, Nov 2002.

[21] X. Shao, J. Yuan, and Y. Shao, "Error performance analysis of linear zero forcing and MMSE precoders for MIMO broadcast channels," *IET Commun.*, vol. 1, no. 5, pp. 1067–1074, Oct 2007.

[22] C.-J. Chen and L.-C. Wang, "Performance analysis of scheduling in multiuser MIMO systems with zero-forcing receivers," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1435–1445, September 2007.

[23] M.-S. Alouini, A. Abdi, and M. Kaveh, "Sum of gamma variates and performance of wireless communication systems over nakagami-fading channels," *IEEE Trans. on Veh. Technol.*, vol. 50, no. 6, pp. 1471–1480, Nov 2001.

[24] J. G. Proakis, *Digital Communications*, 4th ed. New York, NY, USA: McGraw–Hill, 2000.

References

[1] Cisco Visual Networking Index, "Global Mobile Data Traffic Forecast Update, 2013-2018," CISCO, White Paper, Feb. 2014. [Online]. Available: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.pdf

[2] *Measuring the Information Society*. International Telecommunication Union, 2012.

[3] K. Ashton, "That 'internet of things' thing," *RFID Journal*, Jun. 2009. [Online]. Available: http://www.rfidjournal.com/article/view/4986

[4] M. Yun, Y. Rong, Y. Zhou, H.-A. Choi, J.-H. Kim, J. Sohn, and H.-I. Choi, "Analysis of Uplink Traffic Characteristics and Impact on Performance in Mobile Data Networks," in *Proc. of the IEEE International Conference on Communications (ICC)*, Beijing, China, 19–23 May 2008, pp. 4564 –4568.

[5] C. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal*, vol. 27, pp. 379–423 & 623–656, Jul. & Oct. 1948.

[6] G. J. Foschini and M. J. Gans, "On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas," *Wireless Personal Communications*, vol. 6, no. 6, pp. 311–335, 1998.

[7] E. Telatar, "Capacity of Multi-Antenna Gaussian Channels," *European Trans. on Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov. / Dec. 1999.

[8] S. M. Alamouti, "A Simple Transmit Diversity Technique for Wireless Communications," *IEEE J. on Sel. Areas in Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[9] V. Tarokh, N. Seshadri, and A. Calderbank, "Space-Time Codes for High Data Rate Wireless Communication: Performance Criterion and Code Construction," *IEEE Trans. on Inform. Theory*, vol. 44, no. 2, pp. 744–765, 1998.

[10] M. Vu and A. Paulraj, "MIMO Wireless Linear Precoding," *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 86 –105, Sept. 2007.

[11] A. Molisch and M. Win, "MIMO Systems With Antenna Selection," *IEEE Microw. Mag.*, vol. 5, no. 1, pp. 46–56, Mar. 2004.

[12] C. Han, T. Harrold, S. Armour, I. Krikidis, S. Videv, P. Grant, H. Haas, J. Thompson, I. Ku, C.-X. Wang, T. A. Le, M. Nakhai, J. Zhang, and L. Hanzo, "Green Radio: Radio Techniques to Enable Energy-Efficient Wireless Networks," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 46–54, Jun. 2011.

[13] G. Auer, V. Giannini, C. Desset, I. Godor, P. Skillermark, M. Olsson, M. Imran, D. Sabella, M. Gonzalez, O. Blume, and A. Fehske, "How much energy is needed to run a wireless network?" *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 40–49, October 2011.

[14] R. Mesleh, H. Haas, S. Sinanović, C. W. Ahn, and S. Yun, "Spatial Modulation," *IEEE Trans. on Veh. Tech.*, vol. 57, no. 4, pp. 2228 – 2241, Jul. 2008.

[15] M. Di Renzo, H. Haas, and P. M. Grant, "Spatial Modulation for Multiple-Antenna Wireless Systems: A Survey," *IEEE Commun. Mag.*, vol. 49, no. 12, pp. 182–191, December 2011.

[16] M. Di Renzo, H. Haas, A. Ghrayeb, S. Sugiura, and L. Hanzo, "Spatial Modulation for Generalized MIMO: Challenges, Opportunities, and Implementation," *Proc. IEEE*, vol. 102, no. 1, pp. 56–103, Jan 2014.

[17] G. Auer, V. Giannini, C. Desset, I. Godor, P. Skillermark, M. Olsson, M. Imran, D. Sabella, M. Gonzalez, O. Blume, and A. Fehske, "How Much Energy is Needed to Run a Wireless Network?" *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 40–49, 2011.

[18] L. Correia, D. Zeller, O. Blume, D. Ferling, A. Kangas, I. Godor, G. Auer, and L. Van Der Perre, "Challenges and Enabling Technologies for Energy Aware Mobile Radio Networks," *IEEE Commun. Mag.*, vol. 48, no. 11, pp. 66–72, Nov. 2010.

[19] J. Wang, S. Jia, and J. Song, "Generalised Spatial Modulation System with Multiple Active Transmit Antennas and Low Complexity Detection Scheme," *IEEE Trans. on Wireless Commun.*, vol. 11, no. 4, pp. 1605 – 1615, April 2012.

[20] L.-L. Yang, "Transmitter Preprocessing Aided Spatial Modulation for Multiple-Input Multiple-Output Systems," in *Proc. of 73rd IEEE Veh.Techn. Conf. (VTC Spring)*, May 2011, pp. 1 –5.

[21] R. Zhang, L.-L. Yang, and L. Hanzo, "Error Probability and Capacity Analysis of Generalised Pre-coding Aided Spatial Modulation," *IEEE Trans. on Wireless Commun.*, vol. PP, no. 99, pp. 1–1, 2014.

[22] Q. Spencer, A. Swindlehurst, and M. Haardt, "Zero-Forcing Methods for Downlink Spatial Multiplexing in Multiuser MIMO Channels," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 461–471, Feb. 2004.

[23] D. A. Basnayaka, P. J. Smith, and P. A. Martin, "Performance Analysis of Macrodiversity MIMO Systems with MMSE and ZF Receivers in Flat Rayleigh Fading," *IEEE Trans. on Wireless Commun.*, vol. 12, no. 5, pp. 2240–2251, May 2013.

[24] S. Narayanan, M. J. Chaudhary, A. Stavridis, R. Di Renzo, F. Graziosi, and H. Haas, "Multi-User Spatial Modulation MIMO," in *Proc. of IEEE Wireless Communications and Networking Conference (WCNC)*, Apr. 6–9, 2014.

[25] X. Li, Y. Zhang, L. Xiao, X. Xu, and J. Wang, "A Novel Precoding Scheme for Downlink Multi-User Spatial Modulation System," in *Proc. IEEE 24th Int. Symp. Pers. Indoor and Mobile Radio Commun. (PIMRC)*, Sept 2013, pp. 1361–1365.

[26] R. Mesleh, S. Ikki, and M. Alwakeel, "Performance Analysis of Space Shift Keying with Amplify and Forward Relaying," *IEEE Commun. Lett.*, vol. 15, no. 12, pp. 1350–1352, 2011.

[27] P. Yang, B. Zhang, Y. Xiao, B. Dong, S. Li, M. El-Hajjar, and L. Hanzo, "Detect-and-Forward Relaying Aided Cooperative Spatial Modulation for Wireless Networks," *IEEE Trans. on Commun.*, vol. PP, no. 99, pp. 1–12, 2013, (accepted for publication).

[28] G. Auer, V. Giannini, I. Godor, P. Skillermark, M. Olsson, M. Imran, D. Sabella, M. Gonzalez, C. Desset, and O. Blume, "Cellular Energy Efficiency Evaluation Framework," in *Proc. of the Vehicular Technology Conference (VTC Spring)*, May 2011, pp. 1–6.

[29] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time Block Codes from Orthogonal Designs," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1456–1467, Jul. 1999.

[30] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[31] 3GPP, "Further Advancements for E-UTRA Physical Layer Aspects (Release 9)," 3GPP TR 36.814 V0.4.1 (2009-02), Sep. 2009. Retrieved Jun. 2, 2009 from www.3gpp.org/ftp/Specs/.

[32] T. K. Sarkar, R. J. Mailloux, A. A. Oliner, M. Salazar-Palma, and D. L. Sengupta, *History of Wireless*. JOHN WILEY & SONS, INC., 2006.

[33] L. Brandenburg and A. Wyner, "Capacity of the Gaussian Channel with Memory: The Multivariate Case," Bell Systems, Tech. Rep., June 1974.

[34] J. Salz, "Digital Transmission over Cross-Coupled Linear Channels," *AT & T Technical Journal*, vol. 64, pp. 1147–1159, Aug. 1985.

[35] A. J. Paulraj and T. Kailath, "Increasing capacity in wireless broadcast systems using distributed transmission/directional reception," USA Patent 5 345 599, 1994.

[36] E. Larsson, "MIMO Detection Methods: How They Work," *IEEE Signal Process. Mag.*, vol. 26, no. 3, pp. 91 –95, May 2009.

[37] M. Damen, K. Abed-Meraim, and M. Lemdani, "Further Results on the Sphere Decoder," in *Proc. of the IEEE International Symposium on Information Theory*, Washington, DC, Jun. 24–29 2001, pp. 333–.

[38] B. Hassibi and H. Vikalo, "On the Sphere-Decoding Algorithm I. Expected Complexity," *IEEE Trans. on Signal Process.*, vol. 53, no. 8, pp. 2806–2818, Aug. 2005.

[39] L. Barbero and J. Thompson, "Fixing the Complexity of the Sphere Decoder for MIMO Detection," *IEEE Trans. on Wireless Commun.*, vol. 7, no. 6, pp. 2131–2142, Jun. 2008.

[40] J. Jalden, L. Barbero, B. Ottersten, and J. Thompson, "The Error Probability of the Fixed-Complexity Sphere Decoder," *IEEE Trans. on Signal Process.*, vol. 57, no. 7, pp. 2711–2720, July 2009.

[41] A. Klein, G. K. Kaleh, and P. W. Baier, "Zero Forcing and Minimum Mean-Square-Error Equalization for Multiuser Detection in Code-Division Multiple-Access Channels," *IEEE Trans. Veh. Technol.*, vol. 45, no. 2, pp. 276–287, May 1996.

[42] D. Gore, R. Heath, and A. Paulraj, "Transmit Selection in Spatial Multiplexing Systems," *IEEE Commun. Lett.*, vol. 6, no. 11, pp. 491–493, Nov 2002.

[43] J. Mietzner, R. Schober, L. Lampe, W. H. Gerstacker, and P. A. Höeher, "Multiple-Antenna Techniques for Wireless Communications - A Comprehensive Literature Survey," *IEEE Commun. Surveys Tutorials*, vol. 11, no. 2, pp. 87 –105, 2009.

[44] P. Wolniansky, G. Foschini, G. Golden, and R. Valenzuela, "V-BLAST: an Architecture for Realizing very High Data Rates over the Rich-Scattering Wireless Channel," in *Unino Radio-Scientifique Internationale (URSI) Intern. Symp. on Signals, Systems, and Electronics (ISSSE)*, Sep. 29–Oct. 2, 1998, pp. 295–300.

[45] S. Loyka and F. Gagnon, "Performance Analysis of the V-BLAST Algorithm: An Analytical Approach," *EEE Trans. on Wireless Commun.*, vol. 3, no. 4, pp. 1326–1337, July 2004.

[46] H. Zhu, W. Chen, B. Li, and F. Gao, "An Improved Square-Root Algorithm for V-BLAST Based on Efficient Inverse Cholesky Factorization," *IEEE Trans. on Wireless Commun.*, vol. 10, no. 1, pp. 43–48, January 2011.

[47] S. Buzzi, M. Lops, and S. Sardellitti, "Widely Linear Reception Strategies for Layered Space-Time Wireless Communications," *IEEE Trans. on Signal Process.*, vol. 54, no. 6, pp. 2252–2262, June 2006.

[48] D. Basnayaka, P. Smith, and P. Martin, "The Effect of Macrodiversity on the Performance of MLD in Flat Rayleigh/Rician Fading," *IEEE Commun. Lett.*, vol. 16, no. 11, pp. 1764–1767, November 2012.

[49] D. Gesbert, S. Hanly, H. Huang, S. Shamai Shitz, O. Simeone, and W. Yu, "Multi-Cell MIMO Cooperative Networks: A New Look at Interference," *IEEE J. on Sel. Areas in Commun.*, vol. 28, no. 9, pp. 1380–1408, December 2010.

[50] S. Diggavi, N. Al-Dhahir, A. Stamoulis, and A. Calderbank, "Great Expectations: The Value of Spatial Diversity in Wireless Networks," *IEEE Proceedings*, vol. 92, no. 2, pp. 219–270, Feb 2004.

[51] K.-S. Ahn, "Performance Analysis of MIMO-MRC System in the Presence of Multiple Interferers and Noise Over Rayleigh Fading Channels," *IEEE Trans. on Wireless Commun.*, vol. 8, no. 7, pp. 3727–3735, July 2009.

[52] I. Al Falujah and V. Prabhu, "Error Rates of DPSK Systems with MIMO EGC Diversity Reception over Rayleigh Fading Channels," *IEEE Trans. on Commun.*, vol. 56, no. 6, pp. 897–903, June 2008.

[53] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-Time Block Coding for Wireless Communications: Performance Results," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 3, pp. 451–460, Mar. 1999.

[54] T. K. Y. Lo, "Maximum Ratio Transmission," *IEEE Trans. on Commun.*, vol. 47, no. 10, pp. 1458–1461, Oct 1999.

[55] A. Wittneben, "A New Bandwidth Efficient Transmit Antenna Modulation Diversity Scheme for Linear Digital Modulation," in *Proc. of the IEEE International Conference on Communications (ICC)*, vol. 3, May 1993, pp. 1630–1634 vol.3.

[56] M.-O. Damen, K. Abed-Meraim, and J. C. Belfiore, "Diagonal Algebraic Space-Time Block Codes," *IEEE Trans. on Inf. Theory*, vol. 48, no. 3, pp. 628–636, Mar 2002.

[57] Y. Xin, Z. Wang, and G. Giannakis, "Space-Time Diversity Systems Based on Linear Constellation Precoding," *IEEE Trans. on Wireless Commun.*, vol. 2, no. 2, pp. 294–309, Mar 2003.

[58] H. Jafarkhani, "A Quasi-Orthogonal Space-Time Block Code," *IEEE Trans. on Commun.*, vol. 49, no. 1, pp. 1 –4, jan 2001.

[59] B. Hochwald, T. Marzetta, T. Richardson, W. Sweldens, and R. Urbanke, "Systematic Design of Unitary Space-Time Constellations," *IEEE Trans. on Inf. Theory*, vol. 46, no. 6, pp. 1962–1973, Sep 2000.

[60] H. El-Gamal and M.-O. Damen, "Universal Space-Time Coding," *IEEE Trans. on Infor. Theory*, vol. 49, no. 5, pp. 1097–1119, May 2003.

[61] H. Jafarkhani and N. Seshadri, "Super-Orthogonal Space-Time Trellis Codes," *IEEE Trans. on Inf. Theory,*, vol. 49, no. 4, pp. 937–950, April 2003.

[62] Y. Hong, J. Yuan, B. Vucetic, and Z. Chen, "Design of Space-Time Turbo Trellis Codes for Two, Three and Four Transmit Antennas," in *in Proc. the 8th Intern. Conf. on Commun. Systems, 2002 ICCS 2002.*, vol. 1, Nov 2002, pp. 203–207 vol.1.

[63] B. Hassibi and B. Hochwald, "High-Rate Codes That are Linear in Space and Time," *IEEE Trans. on Inf. Theory*, vol. 48, no. 7, pp. 1804 –1824, jul 2002.

[64] L. Zheng and D. N. Tse, "Diversity and Multiplexing: A Fundamental Tradeoff in Multiple-Antenna Channels," *IEEE Trans. on Inform. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[65] V. Tarokh and H. Jafarkhani, "A Differential Detection Scheme for Transmit Diversity," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 7, pp. 1169 –1174, july 2000.

[66] L. Godara, "Applications of Antenna Arrays to Mobile Communications. I. Performance Improvement, Feasibility, and System Considerations," *IEEE Proceedings*, vol. 85, no. 7, pp. 1031–1060, Jul 1997.

[67] T. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. Wong, J. Schulz, M. Samimi, and F. Gutierrez, "Millimeter Wave Mobile Communications for 5G Cellular: It Will Work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.

[68] T. Rappaport, F. Gutierrez, E. Ben-Dor, J. Murdock, Y. Qiao, and J. Tamir, "Broadband Millimeter-Wave Propagation Measurements and Models Using Adaptive-Beam Antennas for Outdoor Urban Cellular Communications," *IEEE Trans. Antennas Propag.*, vol. 61, no. 4, pp. 1850–1859, April 2013.

[69] N. Valliappan, A. Lozano, and R. Heath, "Antenna Subset Modulation for Secure Millimeter-Wave Wireless Communication," *IEEE Trans. on Commun.*, vol. 61, no. 8, pp. 3231–3245, August 2013.

[70] A. Mehana and A. Nosratinia, "Diversity of MIMO Linear Precoding," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 1019–1038, Feb 2014.

[71] M. Shenouda and T. Davidson, "Tomlinson-Harashima Precoding for Broadcast Channels with Uncertainty," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1380–1389, September 2007.

[72] C. Peel, B. Hochwald, and A. Swindlehurst, "A Vector-Perturbation Technique for Near-Capacity Multiantenna Multiuser Communication-Part I: Channel Inversion and Regularization," *IEEE Trans. on Commun.*, vol. 53, no. 1, pp. 195 – 202, Jan. 2005.

[73] A. Stavridis, D. Basnayaka, M. Di Renzo, and H. Haas, "Average Bit Error Probability of Receive-Spatial Modulation Using Zero-Forcing Precoding," in *IEEE 19th Int. Workshop on Computer Aided Modeling and Des. of Commun. Links and Netw. (CAMAD)*, 2014.

[74] A. Wiesel, Y. Eldar, and S. Shamai, "Zero-Forcing Precoding and Generalized Inverses," *IEEE Trans. on Signal Process.*, vol. 56, no. 9, pp. 4409 –4418, Sept. 2008.

[75] P. Viswanath and D. Tse, "Sum Capacity of the Vector Gaussian Broadcast Channel and Uplink-Downlink Duality," *IEEE Trans. on Inform. Theory*, vol. 49, no. 8, pp. 1912–1921, Aug. 2003.

[76] A. del Coso, U. Spagnolini, and C. Ibars, "Cooperative Distributed MIMO Channels In Wireless Sensor Networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 402–414, 2007.

[77] S. Jayaweera, "V-BLAST-Based Virtual MIMO for Distributed Wireless Sensor Networks," *IEEE Trans. on Commun.*, vol. 55, no. 10, pp. 1867–1872, 2007.

[78] H. Chen, A. Gershman, and S. Shahbazpanahi, "Filter-and-Forward Distributed Beamforming in Relay Networks With Frequency Selective Fading," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1251–1262, 2010.

[79] L. Lai, K. Liu, and H. El-Gamal, "The Three-Node Wireless Network: Achievable Rates and Cooperation Strategies," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 805–828, 2006.

[80] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative Strategies and Capacity Theorems for Relay Networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037–3063, 2005.

[81] X. Hong, C.-X. Wang, M. Uysal, X. Ge, and S. Ouyang, "Capacity of Hybrid Cognitive Radio Networks With Distributed VAAs," *IEEE Trans. Veh. Technol.*, vol. 59, no. 7, pp. 3510–3523, 2010.

[82] E. Sengul, E. Akay, and E. Ayanoglu, "Diversity Analysis of Single and Multiple Beamforming," *IEEE Trans. on Commun.*, vol. 54, no. 6, pp. 990 –993, June 2006.

[83] J. Andrews, S. Buzzi, W. Choi, S. Hanly, A. Lozano, A. Soong, and J. Zhang, "What Will 5G Be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, June 2014.

[84] C. Yang, S. Han, X. Hou, and A. Molisch, "How Do We Design CoMP to Achieve Its Promised Potential?" *IEEE Wireless Comm.*, vol. 20, no. 1, pp. 67–74, February 2013.

[85] T. L. Marzetta, "Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas," *IEEE Trans. on Wireless Commun.*, vol. 9, pp. 3590 – 3600, Nov. 2010.

[86] F. Rusek, D. Persson, B. K. Lau, E. Larsson, T. Marzetta, O. Edfors, and F. Tufvesson, "Scaling Up MIMO: Opportunities and Challenges With Very Large Arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40 –60, Jan. 2013.

[87] R. Zhang and C. K. Ho, "MIMO Broadcasting for Simultaneous Wireless Information and Power Transfer," *IEEE Trans. on Wireless Commun.*, vol. 12, no. 5, pp. 1989–2001, May 2013.

[88] J. Rubio and A. Pascual-Iserte, "Energy-Aware Broadcast Multiuser-MIMO Precoder Design with Imperfect Channel and Battery Knowledge," *IEEE Trans. on Wireless Commun.*, vol. 13, no. 6, pp. 3137–3152, June 2014.

[89] A. Stavridis, S. Sinanović, M. D. Renzo., and H. Haas, "Energy Evaluation of Spatial Modulation at a Multi-Antenna Base Station," in *Proc. of the 78th IEEE Veh. Tech. Conf. (VTC)*, Las Vegas, USA, Sep. 2–5, 2013.

[90] H. Holtkamp, G. Auer, S. Bazzi, and H. Haas, "Minimizing Base Station Power Consumption," *IEEE J. Sel. Areas Commun.*, vol. PP, no. 99, pp. 1–10, 2013.

[91] H. Haas, "Wireless Data from Every Light Bulb," TED Website, Aug. 2011. [Online]. Available: http://bit.ly/tedvlc

[92] "'Li-fi' via LED Light Bulb Data Speed Breakthrough," BBC News, Oct. 2013. [Online]. Available: http://www.bbc.co.uk/news/technology-24711935

[93] D. Tsonev, S. Videv, and H. Haas, "Light Fidelity (Li-Fi): Towards All-Optical Networking," in *Proc. SPIE, Broadband Access Commun. Technol. VIII*, vol. 9007, Dec. 18 2013. [Online]. Available: http://dx.doi.org/10.1117/12.2044649

[94] H. Burchardt, N. Serafimovski, D. Tsonev, S. Videv, and H. Haas, "VLC: Beyond Point-to-Point Communication," *IEEE Commun. Mag.*, vol. 52, no. 7, pp. 98–105, July 2014.

[95] M. Safari and M. Uysal, "Relay-assisted Free-Space Optical Communication," *IEEE Trans. on Wireless Commun.*, vol. 7, no. 12, pp. 5441–5449, December 2008.

[96] M. Kashani, M. Rad, M. Safari, and M. Uysal, "All-optical amplify-and-forward relaying system for atmospheric channels," *IEEE Commun. Lett.*, vol. 16, no. 10, pp. 1684–1687, October 2012.

[97] M. Kountouris and J. Andrews, "Downlink SDMA with Limited Feedback in Interference-Limited Wireless Networks," *IEEE Trans. on Wireless Commun.*, vol. 11, no. 8, pp. 2730–2741, August 2012.

[98] A. Stavridis and H. Haas, "Performance Evaluation of Space Modulation Techniques in VLC Systems," in *Proc. of IEEE Int. Conf. on Commun. (ICC) (1st VLCN Workshop)*. London, UK: IEEE, Jun., 8–12 2015.

[99] M. Haenggi, J. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 7, pp. 1029 –1046, september 2009.

[100] J. Andrews, F. Baccelli, and R. Ganti, "A tractable approach to coverage and rate in cellular networks," *IEEE Trans. on Commun.*, vol. 59, no. 11, pp. 3122–3134, Nov. 2011.

[101] A. Stavridis, S. Sinanović, M. D. Renzo, H. Haas, and P. Grant, "An Energy Saving Base Station Employing Spatial Modulation," in *IEEE 17th Int. Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, Sep. 17–19 2012, pp. 231 –235.

[102] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. Heath, "Spatially Sparse Precoding in Millimeter Wave MIMO Systems," *IEEE Trans. on Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, March 2014.

[103] R. Mesleh, H. Haas, C. W. Ahn, and S. Yun, "Spatial Modulation – A New Low Complexity Spectral Efficiency Enhancing Technique," in *IEEE International Conference on Communication and Networking in China (CHINACOM)*, Beijing, China, Oct. 25–27, 2006, pp. 1–5.

[104] J. Jeganathan, A. Ghrayeb, and L. Szczecinski, "Spatial Modulation: Optimal Detection and Performance Analysis," *IEEE Commun. Lett.*, vol. 12, no. 8, pp. 545–547, 2008.

[105] J. Wang, S. Jia, and J. Song, "Signal Vector Based Detection Scheme for Spatial Modulation," *IEEE Commun. Lett.*, vol. 16, no. 1, pp. 19 –21, Jan. 2012.

[106] Q. Tang, Y. Xiao, P. Yang, Q. Yu, and S. Li, "A new low-complexity near-ml detection algorithm for spatial modulation," *IEEE Wireless Commun. Lett.*, vol. 2, no. 1, pp. 90–93, February 2013.

[107] P. Yang, Y. Xiao, L. Li, Q. Tang, and S. Li, "An improved matched-filter based detection algorithm for space-time shift keying systems," *IEEE Signal Process. Lett.*, vol. 19, no. 5, pp. 271–274, May 2012.

[108] Y. A. Chau and S.-H. Yu, "Space Modulation on Wireless Fading Channels," in *proc. of the IEEE Vehicular Technology Conference (VTC Fall 2001)*, vol. 3, 7–11 Oct. 2001, pp. 1668–1671.

[109] H. Haas, E. Costa, and E. Schulz, "Increasing Spectral Efficiency by Data Multiplexing using Antenna Arrays," in *Proc. of the 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, vol. 2, Sep. 2002, pp. 610–613.

[110] S. Song, Y. Yang, Q. Xionq, K. Xie, B.-J. Jeong, and B. Jiao, "A Channel Hopping Technique I: Theoretical Studies on Band Efficiency and Capacity," in *IEEE International Conference on Communications, Circuits and Systems*, vol. 1, Jun. 2004, pp. 229 –233 Vol.1.

[111] J. Jeganathan, A. Ghrayeb, L. Szczecinski, and A. Ceron, "Space Shift Keying Modulation for MIMO Channels," *IEEE Trans. on Wireless Commun.*, vol. 8, no. 7, pp. 3692–3703, Jul. 2009.

[112] J. Jeganathan, A. Ghrayeb, and L. Szczecinski, "Generalized Space Shift Keying Modulation for MIMO Channels," in *Proc. IEEE 19th Intern. Symp. on Personal, Indoor and Mobile Radio Commun. PIMRC 2008*, Cannes, France, 15–18 Sep. 2008, pp. 1–5.

[113] A. Younis, N. Serafimovski, R. Mesleh, and H. Haas, "Generalised Spatial Modulation," in *Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, USA, Nov. 2010.

[114] N. Serafimovski, M. Di Renzo, S. Sinanović, R. Y. Mesleh, and H. Haas, "Fractional Bit Encoded Spatial Modulation (FBE–SM)," *IEEE Commun. Lett.*, vol. 14, no. 5, pp. 429–431, May 2010.

[115] Y. Yang and S. Aissa, "Bit-Padding Information Guided Channel Hopping," *IEEE Commun. Lett.*, vol. 15, no. 2, pp. 163 –165, Feb. 2011.

[116] K. Ntontin, M. Di Renzo, A. Perez-Neira, and C. Verikoukis, "Performance Analysis of Multistream Spatial Modulation With Maximum-Likelihood Detection," in *Proc. of 2013 IEEE Global Commun. Conf. (GLOBECOM)*, Dec 2013, pp. 1590–1594.

[117] R. Mesleh, M. Di Renzo, H. Haas, and P. M. Grant, "Trellis Coded Spatial Modulation," *IEEE Trans. on Wireless Commun.*, vol. 9, no. 7, pp. 2349–2361, Jul. 2010.

[118] E. Basar, U. Aygolu, E. Panayirci, and H. V. Poor, "New Trellis Code Design for Spatial Modulation," *IEEE Trans. on Wireless Commun.*, vol. 10, pp. 2670 – 2680, Aug. 2011.

[119] T. Handte, A. Müller, and J. Speidel, "BER Analysis and Optimization of Generalized Spatial Modulation in Correlated Fading Channels," in *Veh. Tech. Conf. Fall (VTC Fall-2009)*, Anchorage, AK, Sep. 20–23 2009, pp. 1 –5.

[120] M. Di Renzo and H. Haas, "Performance Comparison of Different Spatial Modulation Schemes in Correlated Fading Channels," in *proc. of the 2010 IEEE Intern. Conf. on Commun. (ICC),*, May 2010, pp. 1–6.

[121] E. Basar, U. Aygolu, E. Panayirci, and V. H. Poor, "Space-Time Block Coded Spatial Modulation," *IEEE Trans. on Commun.*, vol. 59, no. 3, pp. 823 –832, Mar. 2011.

[122] M. Di Renzo and H. Haas, "On Transmit Diversity for Spatial Modulation MIMO: Impact of Spatial Constellation Diagram and Shaping Filters at the Transmitter," *IEEE Trans. on Veh. Technol.*, vol. 62, no. 6, pp. 2507–2531, July 2013.

[123] ——, "Space Shift Keying MIMO over Correlated Rician Fading Channels: Performance Analysis And a New Method for Transmit-Diversity," *IEEE Trans. on Commun.*, vol. 59, no. 1, pp. 116–129, January 2011.

[124] M.-T. Le, V.-D. Ngo, H.-A. Mai, X. N. Tran, and M. Di Renzo, "Spatially Modulated Orthogonal Space-Time Block Codes with Non-Vanishing Determinants," *IEEE Trans. on Commun.*, vol. 62, no. 1, pp. 85–99, January 2014.

[125] Z. Zhou, N. Ge, and X. Lin, "Reduced-Complexity Antenna Selection Schemes in Spatial Modulation," *IEEE Commun. Lett.*, vol. 18, no. 1, pp. 14–17, January 2014.

[126] R. Rajashekar, K. Hari, and L. Hanzo, "Antenna Selection in Spatial Modulation Systems," *IEEE Commun. Lett.*, vol. 17, no. 3, pp. 521–524, March 2013.

[127] K. Ntontin, M. Di Renzo, A. Perez-Neira, and C. Verikoukis, "A Low-Complexity Method for Antenna Selection in Spatial Modulation Systems," *IEEE Commun. Lett.*, vol. 17, no. 12, pp. 2312–2315, December 2013.

[128] N. Wang, W. Liu, H. Men, M. Jin, and H. Xu, "Further Complexity Reduction Using Rotational Symmetry for EDAS in Spatial Modulation," *IEEE Commun. Lett.*, vol. PP, no. 99, pp. 1–1, 2014, (Accepted for Publication).

[129] S. Sugiura, S. Chen, and L. Hanzo, "Generalized Space-Time Shift Keying Designed for Flexible Diversity-, Multiplexing- and Complexity-Tradeoffs," *IEEE Trans. on Wireless Commun.*, vol. 10, no. 4, pp. 1144 –1153, Apr. 2011.

[130] ——, "Coherent and Differential Space-Time Shift Keying: A Dispersion Matrix Approach," *IEEE Trans. on Commun.*, vol. 58, no. 11, pp. 3219–3230, November 2010.

[131] M. Di Renzo and H. Haas, "Improving the Performance of Space Shift Keying (SSK) Modulation Via Opportunistic Power Allocation," *IEEE Commun. Lett.*, vol. 14, no. 6, pp. 500–502, June 2010.

[132] M. Maleki, H. Bahrami, S. Beygi, M. Kafashan, and N. Tran, "Space modulation with csi: Constellation design and performance evaluation," *IEEE Trans. on Vehicular Techn.*, vol. 62, no. 4, pp. 1623–1634, May 2013.

[133] X. Wu, S. Sinanovic, M. Di Renzo, and H. Haas, "Structure optimisation of spatial modulation over correlated fading channels," in *Proc. of 2012 IEEE Global Commun. Conf. (GLOBECOM)*, Dec 2012, pp. 4049–4053.

[134] X. Wu, M. Di Renzo, and H. Haas, "Direct Transmit Antenna Selection for Transmit Optimized Spatial Modulation," in *Proc. of 78th IEEE Veh.Techn. Conf. (VTC Spring)*, Sept 2013, pp. 1–5.

[135] P. Yang, Y. Xiao, Y. Yu, and S. Li, "Adaptive Spatial Modulation for Wireless MIMO Transmission Systems," *IEEE Commun. Lett.*, vol. 15, no. 6, pp. 602–604, June 2011.

[136] P. Yang, Y. Xiao, Y. Yu, L. Li, Q. Tang, and S. Li, "Simplified Adaptive Spatial Modulation for Limited-Feedback MIMO Systems," *IEEE Trans. on Veh. Tech.*, vol. 62, no. 6, pp. 2656–2666, July 2013.

[137] A. Stavridis, S. Sinanović, M. Di Renzo, and H. Haas, "Transmit Precoding for Receive Spatial Modulation Using Imperfect Channel Knowledge," in *Proc. of 75th IEEE Veh. Techn. Conf. (VTC Spring)*, 2012.

[138] S. Sugiura, C. Sheng, and L. Hanzo, "Generalized Space-Time Shift Keying Designed for Flexible Diversity-, Multiplexing- and Complexity-Tradeoffs," *IEEE Trans. Wireless Commun.*, vol. 10, pp. 1144 – 1153, 2011.

[139] S. Sugiura, "Dispersion Matrix Optimization for Space-Time Shift Keying," *IEEE Commun. Lett.*, vol. 15, no. 11, pp. 1152–1155, November 2011.

[140] R. Rajashekar, K. Hari, and L. Hanzo, "Structured dispersion matrices from division algebra codes for space-time shift keying," *IEEE Signal Process. Lett.*, vol. 20, no. 4, pp. 371–374, April 2013.

[141] S. Sugiura and L. Hanzo, "On the Joint Optimization of Dispersion Matrices and Constellations for Near-Capacity Irregular Precoded Space-Time Shift Keying," *IEEE Trans. on Wireless Commun.*, vol. 12, no. 1, pp. 380–387, January 2013.

[142] M. Di Renzo and H. Haas, "Bit Error Probability of Space-Shift Keying MIMO Over Multiple-Access Independent Fading Channels," *IEEE Trans on Veh. Technol.*, vol. 60, no. 8, pp. 3694–3711, Oct 2011.

[143] N. Serafimovski, S. Sinanovic, M. Renzo, and H. Haas, "Multiple Access Spatial Modulation," *EURASIP J. on Wireless Commun. and Netw.*, vol. 2012, no. 1, pp. 1–20, 2012.

[144] N. Serafimovski, S. Sinanovic, A. Younis, M. Di Renzo, and H. Haas, "2-user multiple access spatial modulation," in *IEEE GLOBECOM Workshops (GC Wkshps)*, Houston, USA, Dec. 5–9 2011, pp. 343–347.

[145] M. Raed, I. Salama, A. Hadi, and A. Mansour, "Performance Analysis of Space Shift Keying (SSK) Modulation with Multiple Cooperative Relays," *EURASIP J. on Adv. in Signal Process.*, 2012.

[146] M. Wen, X. Cheng, V. Poor, and B. Jiao, "Use of Space Shift Keying (SSK) Modulation in Two-Way Amplify-and-Forward Relaying," *IEEE Trans. Veh. Technol.*, vol. PP, no. 99, pp. 1–1, 2013, (early access article).

[147] P. Som and A. Chockalingam, "End-to-End BER Analysis of Space Shift Keying in Decode-and-Forward Cooperative Relaying," in *IEEE Wireless Commun. and Networking Conf. (WCNC)*, 2013, pp. 3465–3470.

[148] ——, "BER Analysis of Space Shift Keying in Cooperative Multi-hop Multi-branch DF Relaying," in *Proc. of the 78th IEEE Veh. Tech. Conf. (VTC)*, Las Vegas, USA, Sep. 2–5, 2013.

[149] S. Sugiura, S. Chen, H. Haas, P. Grant, and L. Hanzo, "Coherent Versus Non-Coherent Decode-and-Forward Relaying Aided Cooperative Space-Time Shift Keying," *IEEE Trans. Commun.*, vol. 59, no. 6, pp. 1707–1719, Jun. 2011.

[150] N. Serafimovski, S. Sinanovic, M. Di Renzo, and H. Haas, "Dual-hop Spatial Modulation (Dh-SM)," in *Proc. of the IEEE Veh. Tech. Conf. (VTC Spring)*, Budapest, Hungary, May 15–18, 2011, pp. 1–5.

[151] S. Narayanan, M. Di Renzo, F. Graziosi, and H. Haas, "Distributed Spatial Modulation for Relay Networks," in *Proc. of the 78th IEEE Veh. Tech. Conf. (VTC)*, Las Vegas, USA, Sep. 2–5, 2013.

[152] ——, "Distributed Space Shift Keying For The Uplink of Relay-Aided Cellular Networks," in *IEEE 17th Int. Workshop on Computer Aided Modeling and Des. of Commun. Links and Netw. (CAMAD)*, 2012, pp. 130–134.

[153] Y. Yang and S. Aissa, "Information-Guided Transmission in Decode-and-Forward Relaying Systems: Spatial Exploitation and Throughput Enhancement," *IEEE Trans. on Wireless Commun.*, vol. 10, no. 7, pp. 2341–2351, 2011.

[154] S. Narayanan, A. Stavridis, M. Di Renzo, F. Graziosi, and H. Haas, "Distributed Spatially-Modulated Space-Time-Block-Codes," in *IEEE 18th Int. Workshop on Computer Aided Modeling and Des. of Commun. Links and Netw. (CAMAD)*, 2013.

[155] A. Stavridis, S. Sinanović, M. D. Renzo, and H. Haas, "A Power Saving Dual-Hop Architecture Based on Hybrid Spatial Modulation," in *2012 Conf. Record of the Forty Sixth Asilomar Conf. on Signals, Systems and Computers (ASILOMAR)*, Nov. 4–7 2012, pp. 1366–1370.

[156] A. Younis, W. Thompson, M. D. Renzo, C.-X. Wang, M. A. Beach, H. Haas, and P. M. Grant, "Performance of Spatial Modulation using Measured Real-World Channels," in *Proc. of the 78th IEEE Veh. Tech. Conf. (VTC)*, Las Vegas, USA, Sep. 2–5 2013.

[157] S. Sinanović, M. D. Renzo, and H. Haas, "Secrecy Rate of Time Switched Transmit Diversity System," in *Proc. of the Vehicular Technology Conference (VTC)*, IEEE. Budapest, Hungary: IEEE, May 15–18, 2011, pp. 1–5.

[158] J. Zhang, Y. Wang, L. Ding, and N. Zhang, "Bit Error Probability of Spatial Modulation over Measured Indoor Channels," *IEEE Trans. on Wireless Commun.*, vol. 13, no. 3, pp. 1380–1387, March 2014.

[159] N. Serafimovski, A. Younis, R. Mesleh, P. Chambers, M. D. Renzo, C.-X. Wang, P. M. Grant, M. A. Beach, and H. Haas, "Practical Implementation of Spatial Modulation," *IEEE Trans. on Veh. Tech.*, vol. 62, no. 9, pp. 4511–4523, 2013.

[160] A. Bansal, M. Bhatnagar, and A. Hjorungnes, "Decoding and Performance Bound of Demodulate-and-Forward Based Distributed Alamouti STBC," *IEEE Trans. on Wireless Commun.*, vol. 12, no. 2, pp. 702–713, February 2013.

[161] Y. Yang and B. Jiao, "Information-Guided Channel-Hopping for High Data Rate Wireless Communication," *IEEE Communications Letters*, vol. 12, no. 4, pp. 225 –227, Apr. 2008.

[162] R. Rajashekar, K. Hari, and L. Hanzo, "Reduced-Complexity ML Detection and Capacity-Optimized Training for Spatial Modulation Systems," *IEEE Trans. on Commun.*, vol. 62, no. 1, pp. 112–125, January 2014.

[163] X. Guan, Y. Cai, and W. Yang, "On the Mutual Information and Precoding for Spatial Modulation with Finite Alphabet," *IEEE Wireless Commun. Lett.*, vol. 2, no. 4, pp. 383–386, August 2013.

[164] P. Frenger, P. Moberg, J. Malmodin, Y. Jading, and I. Godor, "Reducing Energy Consumption in LTE with Cell DTX," in *in Proc. of the IEEE 73rd Veh. Techn. Conf. (VTC Spring)*, May 2011, pp. 1–5.

[165] G. Miao, N. Himayat, and G. Li, "Energy-Efficient Link Adaptation in Frequency-Selective Channels," *IEEE Trans. on Commun.*, vol. 58, no. 2, pp. 545–554, February 2010.

[166] A. Fehske, P. Marsch, and G. Fettweis, "Bit per Joule Efficiency of Cooperating Base Stations in Cellular Networks," in *in Proc of the 2010 IEEE GLOBECOM Workshops (GC Wkshps)*, Dec 2010, pp. 1406–1411.

[167] S. Sandhu and A. Paulraj, "Space-Time Block Codes: A Capacity Perspective," *IEEE Commun. Lett.*, vol. 4, no. 12, pp. 384–386, 2000.

[168] A. Stavridis, S. Narayanan, M. Di Renzo, L. Alonso, H. Haas, and C. Verikoukis, "A Base Station Switching on-off Algorithm Using Traditional MIMO and Spatial Modulation," in *IEEE 18th Int. Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, Sept 2013, pp. 68–72.

[169] J. G. Proakis and M. Salehi, *Communication System Engineering*. Prentice Hall, 1994.

[170] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed., A. Jeffrey and D. Zwillinger, Eds. Academic Press, Mar. 2007, ISBN-10: 0123736374.

[171] N. Balakrishnan and C. D. Lai, *Continuous Bivariate Distributions*. Springer, 2009.

[172] M. Chiani, D. Dardari, and M. K. Simon, "New Exponential Bounds and Approximations for the Computation of Error Probability in Fading Channels," *IEEE Trans. on Wireless Commun.*, vol. 2, no. 4, pp. 840–845, July 2003.

[173] Z. Wang and G. Giannakis, "A Simple and General Parameterization Quantifying Performance in Fading Channels," *IEEE Trans. on Commun.*, vol. 51, no. 8, pp. 1389–1398, Aug 2003.

[174] W. Ho, T. Quek, S. Sun, and R. Heath, "Decentralized Precoding for Multicell MIMO Downlink," *IEEE Trans. on Wireless Commun.*, vol. 10, no. 6, pp. 1798 –1809, June 2011.

[175] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[176] A. Tulino, A. Lozano, and S. Verdu, "Impact of Antenna Correlation on the Capacity of Multiantenna Channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2491 – 2509, July 2005.

[177] A. M. Tulino and S. Verdu, "Random Matrix Theory and Wireless Communications," *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 1, pp. 1–182, 2004.

[178] G. Alfano, A. Tulino, A. Lozano, and S. Verdu, "Capacity of MIMO Channels With One-Sided Correlation," in *Proc. of 2004 IEEE Eighth International Symposium on Spread Spectrum Techniques and Applications*, Aug 2004, pp. 515–519.

[179] A. Stavridis, D. Basnayaka, S. Sinanovic, M. Di Renzo, and H. Haas, "A Virtual MIMO Dual-Hop Architecture Based on Hybrid Spatial Modulation," *IEEE Trans. on Commun.*, vol. 62, no. 9, pp. 3161–3179, Sept 2014.

[180] X. Shao, J. Yuan, and Y. Shao, "Error Performance Analysis of Linear Zero Forcing and MMSE Precoders for MIMO Broadcast Channels," *IET Commun.*, vol. 1, no. 5, pp. 1067–1074, Oct 2007.

[181] C.-J. Chen and L.-C. Wang, "Performance Analysis of Scheduling in Multiuser MIMO Systems with Zero-Forcing Receivers," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1435–1445, September 2007.

[182] M.-S. Alouini, A. Abdi, and M. Kaveh, "Sum of Gamma Variates and Performance of Wireless Communication Systems Over Nakagami-Fading Channels," *IEEE Trans. on Veh. Technol.*, vol. 50, no. 6, pp. 1471–1480, Nov 2001.

[183] J. G. Proakis, *Digital Communications*, 4th ed. New York, NY, USA: McGraw–Hill, 2000.

[184] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. McGraw–Hill, 1991.

[185] N. Jindal and A. Goldsmith, "Dirty-paper coding versus TDMA for MIMO Broadcast channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 5, pp. 1783–1794, May 2005.

[186] M. O. Hasna and M.-S. Alouini, "End-to-end Performance of Transmission Systems with Relays over Rayleigh-fading Channels," *IEEE Transactions on Wireless Communications*, vol. 2, no. 6, pp. 1126 – 1131, Nov. 2003.

[187] A. Younis, D. A. Basnayaka, and H. Haas, "Performance Analysis for Generalised Spatial Modulation," in *Proc. of European Wireless Conf. (EW 2014)*, Barcelona, Spain, 14–16 May 2014, pp. 207–212.

[188] P. Tejera, W. Utschick, J. Nossek, and G. Bauch, "Rate Balancing in Multiuser MIMO OFDM Systems," *IEEE Trans. on Commun.*, vol. 57, no. 5, pp. 1370–1380, May 2009.

[189] A. Stavridis, M. Di Renzo, and H. Haas, "Performance Analysis of Multi-Stream Receive Spatial Modulation in the MIMO Broadcast Channel," *IEEE Trans. on Wireless Commun.*, 2015, (Accepted for Publication).

[190] Muhammad Ijaz, Dobroslav Tsonev, Athanasios Stavridis, Abdelhamid Younis Jonathan J. D. McKendry, Erdan Gu, Martin D. Dawson, Stefan Videv, and Harald Haas, "Optical Spatial Modulation OFDM using Micro-LEDs," in *Asilomar Conference on Signals, Systems, and Computersns, Pacific Grove, CA, USA*, Nov 2014.

[191] A. Stavridis, M. Renzo, and H. Haas, "On the Performance of Multi-Stream Receive Spatial Modulation in the MIMO Broadcast Channel," in *Proc. of IEEE Global Commun. Conf. (GLOBECOM)*. San Diego, CA, USA: IEEE, Dec., 6–10 2015, (to appear).